ARL TECHNICAL REPORT 60-275

ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS

Karl G. Guderley
Mary D. Lum

FEBRUARY 1961

AERONAUTICAL RESEARCH LABORATORY
AIR FORCE RESEARCH DIVISION

NOTICES

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Qualified requesters may obtain copies of this report from the Armed Services Technical Information Agency, (ASTIA), Arlington Hall Station, Arlington 12, Virginia.

This report has been released to the Offices of Technical Services, U. S. Department of Commerce, Washington 25, D. C. for sale to the general public.

Copies of ARL Technical Reports and Technical Notes should not be returned to Aeronautical Research Laboratories unless return is required by security considerations, contractual obligations, or notices on a specific document.

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)
Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities
I. Guderley, K.G.
II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)
Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities
I. Guderley, K.G.
II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)
Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities

I. Guderley, K.G.
II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)
Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities

I. Guderley, K.G.
II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

ARL TECHNICAL REPORT 60-275

ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS

Karl G. Guderley
Mary D. Lum

FEBRUARY 1961

Project No. 7071
Task 70437

AERONAUTICAL RESEARCH LABORATORY
AIR FORCE RESEARCH DIVISION
AIR RESEARCH AND DEVELOPMENT COMMAND
UNITED STATES AIR FORCE
WRIGHT-PATTERSON AIR FORCE BASE, OHIO

600 – April 1961 – 26-960

FOREWORD

This report was prepared in the Applied Mathematics Research Branch of the Aeronautical Research Laboratory, Air Force Research Division, Air Research and Development Command, Wright-Patterson Air Force Base, Ohio. The problem stems from a proposed electronic method for extracting information from very low-contrast photographs, through grain counts of the enlargements, and originated in the Solid State Physics Research Branch of the Aeronautical Research Laboratory. The authors gratefully acknowledge numerous discussions with Dr. Lee Devol and Mr. Radames Gebel on the basic physical question which has been presented in WADC TN 58-110, "Electronic Contrast Selector and Grain Spacing to Light Intensity Translator for Photographic Enlargements", by R. Gebel. The work presented herein is an attempt at a suitable mathematical formulation as a basis for proper analysis of such grain counts. The authors also acknowledge the assistance of Mrs. Martha Elmore and Mr. James Caslin in the preparation of some of the tables and graphs. Appendix III on computational procedures developed for formulas given in Section 7 (to be used with the Burroughs E-101 Electrodata Electronic Computer) was written by Mr. Caslin. The work was carried out under Project 7071, Task 70437, "Methods of Mathematical Physics".

ABSTRACT

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test are") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches a weight factor to each grain depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one can also scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors in the measurement process: errors due to the non-vanishing size of the test area, and errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized. The treatment of these questions is shown in a number of examples of increasing complexity. Moreover, this report investigates

the probability of making an error if one tries to discriminate between two known light densities on the basis of grain counts, and it also examines a step-wise method for carrying out such counts.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Continued)

LIST OF TABLES

INTRODUCTION

In order to gain maximum information from a photograph one may enlarge it to the extent that the individual grains show up. The density of the grains is proportional to the flux of light hitting a particular area. The grains of the photographic plate arise by a random process; therefore, the statistical variations in the density of them set a limit to the evaluation. The interpretation of the photograph will then amount to somehow forming average densities; also one may try to recognize patterns or some main features of the picture. This process is usally done by inspection, perhaps with some additional photographic techniques, such as high contrast prints. The question can be asked whether this process can be done by a photoelectric scanning process. This may be desirable for several reasons. The photoelectric processes might reveal information which is hard to recognize with the eye. Whether this is true depends upon the performance of the eye, and this question does not concern us in this report. The photoelectric evaluation would not require human judgment, and thus it is not subject to fatigue and also it can be done more quickly. It is felt that the second point is sufficiently important to justify a closer study of this possibility.

We ask in this report which conclusion about the energy density of the electromagnetic waves can be drawn from the distribution of grains on given photographic plates. Local measurements are impossible because of the grain structure of the plate, therefore one will try to characterize the desired light density at a given point by an average - possibly a weighted average - of the grain density over an area surrounding this point. This paper treats this process from a probabilistic point of view. Confidence limits are established for different measuring procedures and it is shown how to choose a weight factor characterizing the measurements in such a manner that the error limits are minimized.

By a suitable choice of the weight factor it is possible to determine, beside the light density, other quantities related to it, e.g. its gradient. Furthermore, it is possible to search for special patterns, e.g. for a jump of the light density, or for narrow lines of a higher density.

From a technical point of view the question of discriminating between two fixed light densities may be of interest. Here one will define a certain cut-off count for the number of grains, and ascribe to all areas with a count below the cut-off the lower light density, to all others the higher one. For this procedure, formulae for the probability of error are given.

From a practical point of view, one might be inclined to determine first for subareas of the test area, whether they belong to the higher or to the lower light density. The majority of assignment of one kind or the other in the test area will then determine, whether the test area belongs to the higher or to the lower count. The probability of error for this procedure and for a direct count for the entire test area is also investigated.

1. Model Connecting the Grains of the Photographic Plate With the Light Flux

For our theoretical analysis we need a model for the interaction of light with the photographic plate. The light manifests itself in photons, i. e. at discrete points of the photographic plate. Not every photon gives a chemical reaction on the plate; a major portion of the light passes through without indication. But some photons modify certain molecules. The grains which we see are clusters of changed molecules which grow because of the development process around the molecules of the photographic emulsion that have been changed by the photons. The subsequent computations are based on the following model for the creation of grains: For a first example assume the energy density of the light waves to be constant. Let us assume that we can stretch the time scale in such a manner that one grain arises after the other. Then the grain arising first will be found at some spot of the plate and all locations are equally probable. The probability of finding a grain in a given area is equal to this area divided by the area of the whole plate. The second grain arises in exactly the same manner and its location is not influenced by the location of the first grain. The same holds for all grains that arise subsequently. One says that the arising of grains represent independent events.

If the light density varies across the plate, then the probability of finding the first grain in a given area is equal to the integral of the light flux for the area considered, divided by the integral for the light flux over the whole plate. Again the production of any one grain is assumed to be independent of the production of all other grains. Furthermore it is assumed that all grains have the same darkness and can be considered as points, i.e. if one counts grains in a given area, a grain is either inside or outside of that area.

2. Some Results of Probability Theory

For our future discussion some results of probability theory will be needed. They are derived in Appendix I. Here they will be quoted and described in a self-contained manner.

The number of photons that act on the photographic plate and the number of grains produced are considered as very large. For the measurement, a small area called "test area" of the photographic plate is considered and from the number and the arrangement of the grains certain conclusions on the density of the incoming light are drawn.

Let us introduce on the phographic plate a Cartesian system of coordinates. The probability that a grain is produced in a certain area is considered to be proportional to the light flux across the area multiplied by the exposure time. Let the density of the light flux (the energy of the electromagnetic waves per unit area) be given by $\rho(x, y)$. If one makes a great number of photographs of the same object and counts in all photographs the number of grains in the same test area A, then the counts will in general not be the same. Let N be the conceptual average formed from an infinitely great number of such counts. (The number N is also denoted as "expected value"). A factor of proportionality which contains the exposure time and the sensitivity of the plate may be denoted by k such that

$$N = k \iint_A \rho(x, y)\, dxdy \tag{1}$$

is the average number of grains to be expected in A.

Let $\theta(x, y)$ be a weighting function. The choice of this weighting function determines the measurement that is carried out. The individual grains found in the test area in a specific experiment may be numbered as 1, 2, ... i ... n (n will be close to N but need not be equal to N); the coordinates of the individual grains are (x_1, y_1) ... (x_i, y_i) ... The function $\theta(x, y)$ then yields the weight $\theta(x_i, y_i)$ for grain i in the position (x_i, y_i). A measurement then determines the expression

$$Y = \Sigma\, \theta(x_i, y_i) \tag{2}$$

where the summation is to be extended over all points that lie in the test area. For a fixed area A and a given photographic plate, Y will have a fixed value. If one evaluates a great number of plates which have arisen from the same experiment, i. e. from the same function $\rho(x, y)$ with the same constant k, one will obtain different values of

Y. Let p(u) be the probability density of Y taking on values u. Fig. 1 shows such a probability density curve. The area under the curve is the probability of finding a test result between $u = -\infty$ and $u = +\infty$. This probability is 1. The average value of Y is defined by

$$M_Y = \int_{-\infty}^{+\infty} u \; p(u) \, du \tag{3}$$

In our measurements the quantity that is to be determined will be approximately proportional to the average value; but naturally, individual measurements will give results that are scattered around the average value. To characterize the width of the scatter one introduces the variance which is defined by

$$V_Y = \int_{-\infty}^{+\infty} (u - M_Y)^2 \; p(u) \, du \tag{4}$$

To show how the variance is connected with the width of the curve, let us first assume that the probability is constant over values Y which deviate from the average by at most b. (In other words b is the half width of a rectangular distribution). Then the probability density is 1/2b and one obtains for the variance

$$V_Y = \frac{1}{2b} \int_{-b}^{+b} (u - M_Y)^2 \, d(u - M_Y) = \frac{b^2}{3}$$

i.e. in this case

$$b = \tau \sqrt{V_Y} \tag{5}$$

where $\tau = \sqrt{3}$

Frequently, e.g. if the number of points in the test area is large, the distribution will be close to normal, as can be shown by means of the Central Limit Theorem (see [1] page 180). Then, since the distribution extends to infinity, b can no longer be defined as the half width of the area which contains all the points. It is then considered as the half width of an area that contains a fraction $(1 - \alpha)$ of all points. We can use (5) again; τ is given as a function of α in the following table

α	$1 - \alpha$	τ
.317	.683	1
.045	.955	2
.003	.997	3
.100	.900	1.64
.050	.950	1.96
.010	.990	2.58

The value of τ which we obtained previously for a rectangular distribution then corresponds to an α of about 0.1, i.e. 90% of the measurements will fall within the region

$$M_Y \pm 1.6 \sqrt{V_Y}$$

It is not impossible to determine the probability density curves for the experiments to be described in this report. They may be obtained from the characteristic functions given in Appendix I. Here we mention only two results. The average value of infinitely many measurements (i.e. for an identical test area in infinitely many photographic plates) is given by

$$M_Y = k \iint_A \rho(x, y)\, \theta(x, y)\, dxdy \tag{6}$$

The variance is given by

$$V_Y = k \iint_A \rho(x, y)\, \theta^2(x, y)\, dxdy \tag{7}$$

In making measurements one will define θ in such a manner, that the average value M_Y is directly related to the quantity which one wants to find. Then one can compute from (7) the variance incurred in such a measurement and determine from it probable bounds for the errors which may arise, if one uses Y as an approximation for M_Y. Examples will be given in the next section.

Later another result of probability theory will be needed. The quantity as defined by the right hand side of (6) is in general only an approximation to the quantity which is to be determined. One reason is, that the form of (6) is an average over an area, while we are frequently interested in local values. To fix the ideas take the local value of ρ (x, y). This local value of ρ must be approximated by an integral M_Y over an area. In general this average M_Y will not be equal to the value of ρ at the desired point. The difference depends upon the character of the functions ρ and θ, and it will vary with x and y; basically it depends upon the properties of the photograph. Thus the quantity M_Y - even if suitably defined - is only an approximation to ρ and the difference $M_Y - \rho$ possesses a variance, which we shall denote by V_1. In estimating M_Y by counting the number of grains within the area and calculating Y another error is incurred which has a variance V_Y. Now probability theory gives the following result: If the two probabilistic processes are independent then the result will have a variance given by the sum of the two variances.

3. Examples

Some examples may show how these results are applied to specific situations. (a) Let us assume that ρ is constant (but unknown) and that we want to determine its value. Taking a rectangular test area with sides 2a and 2b one finds

$$M_Y = k\rho \int_{-a}^{+a} \int_{-b}^{+b} \theta(x, y)\, dxdy$$

$$\rho = \frac{M_Y}{k \int_{-a}^{a} \int_{-b}^{b} \theta(x, y)\, dxdy}$$

The last equation shows the relation between the desired value ρ and the average value M_Y. ρ is proportional to M_Y and the factor of proportionality is known. The variance is given by

$$V_Y = k\rho \int_{-a}^{+a} \int_{-b}^{+b} \theta^2 (x, y)\, dxdy$$

In these formulae θ is still arbitrary. Let

$$\theta \equiv 1$$

then

$$M_Y = k\rho 4ab$$

M_Y is then the average number of points N in the area. The variance is

$$V_Y = k\rho 4ab = M_Y = N$$

Using Eq. (5) for the width of the probability curve one finds as measure for the relative error

$$\epsilon = \frac{b}{M_Y} = \frac{\tau}{\sqrt{N}} \qquad (8)$$

Here τ must be taken from the table given in Section 2. If $\alpha = .003$ and $N = 10,000$ one would obtain a relative error in the order of 3%; for $N = 100$ one would obtain 30% relative error. For $\alpha = .10$ the relative error with 1000 points is 5.2%. This may be considered as tolerable. About 1000 points in the test area appears to be the minimum that should be used.

One might try to reduce the variance by choosing another function θ. Unfortunately $\theta \equiv$ constant is the best choice. Let us pose the problem of minimizing the variance while the average value M_Y is kept constant:

$$\iint_A \theta(x, y)\,dxdy = \text{constant}$$

$$\iint_A \theta^2(x, y)\,dxdy = \text{minimum}$$

This is a simple problem of the calculus of variations. Introducing a Lagrangian multiplier λ one obtains an equivalent problem:

$$\iint_A \{\theta^2(x, y) + \lambda\,\theta(x, y)\}\,dxdy = \text{minimum}$$

Hence

$$\iint_A \delta\,\theta\,\{2\,\theta(x, y) + \lambda\}\,dxdy = 0$$

or

$$\theta = -\frac{\lambda}{2} = \text{constant}$$

In computing the relative error the choice of the constant in the last equation is unessential. We set $\theta = 1$, and thus find that the above choice of θ was optimal.

(b) Let

$$\rho = \rho_0 [1 + a_1 \frac{x}{a} + b_1 \frac{y}{b}] ; \tag{9}$$

we try to determine ρ at $x = 0$, $y = 0$. Choosing $\theta = 1$ we obtain

$$M_Y = k \rho_0 4ab$$

Thus M_Y is proportional to ρ_0 at the origin, which shows that this choice of θ is suitable for the desired measurement. The variance is found to be

$$V_Y = k \rho_0 \int_{-a}^{+a} \int_{-b}^{+b} (1 + a_1 \frac{x}{a} + b_1 \frac{y}{b}) \, dxdy$$

$$= 4 k \rho_0 ab$$

Again these quantities can be expressed by means of the average number of points N in the test area, and one obtains the same result for the relative error as in (8.)

Let us again consider the minimum problem for θ. The function $\theta(x, y)$ must be chosen such that only the constant term in the above expression for ρ will give a contribution to M_Y. Let

$$\rho_1(x, y) = 1$$

$$\rho_2(x, y) = x/a$$

$$\rho_3(x, y) = y/b$$

Then we must determine $\theta(x, y)$ in such a manner that

$$\int_{-a}^{+a}\int_{-b}^{+b} \theta(x, y)\, \rho_1(x, y)\, dxdy = 1 \tag{10a}$$

$$\int_{-a}^{+a}\int_{-b}^{+b} \theta(x, y)\, \rho_2(x, y)\, dxdy = 0 \tag{10b}$$

$$\int_{-a}^{+a}\int_{-b}^{+b} \theta(x, y)\, \rho_3(x, y)\, dxdy = 0 \tag{10c}$$

$$\int_{-a}^{+a}\int_{-b}^{+b} (\rho_1 + a_1\, \rho_2 + b_1\, \rho_3)\, \theta^2(x, y) dxdy = \text{minimum} \tag{10d}$$

By variational calculus one obtains

$$\theta(x, y) = \frac{\lambda_1\, \rho_1 + \lambda_2\, \rho_2 + \lambda_3\, \rho_3}{\rho_1 + a_1\, \rho_2 + b_1\, \rho_3}$$

where λ_1, λ_2 and λ_3 must be determined from conditions (10a, b, c).

By chance this problem possesses a simple solution. If we set

$$\lambda_1 = 1$$

$$\lambda_2 = a_1$$

$$\lambda_3 = b_1$$

one finds $\theta \equiv 1$. The conditions (10b) and (10c) are automatically fulfilled. The condition (10a) is unessential, it could be fulfilled by choosing a suitable constant for θ, instead of $\theta = 1$. But such a constant would drop out in the determination of the relative accuracy ϵ.

Thus we find, that even for a ρ surface that possesses a gradient, $\theta \equiv 1$ is the best choice for the weight function.

(c) Going one step further one might consider a function ρ which is represented in a certain area by

$$\rho = \rho_0 \,(1 + a_1 x + a_2 y + a_3 x^2 + a_4 xy + a_5 y^2)$$

To obtain sumple results we assume that ρ is nearly constant. Then by calculus of variations one finds

$$\theta \sim \lambda_1 + \lambda_2 x + \lambda_3 y + \lambda_4 x^2 + \lambda_5 xy + \lambda_6 y^2$$

with the additional conditions

$$\iint_A \theta \, dxdy = 1 \qquad (11a) \qquad\qquad \iint_A \theta x^2 \, dxdy = 0 \qquad (11d)$$

$$\iint_A \theta x \, dxdy = 0 \qquad (11b) \qquad\qquad \iint \theta xy \, dxdy = 0 \qquad (11e)$$

$$\iint_A \theta y \, dxdy = 0 \qquad (11c) \qquad\qquad \iint \theta y^2 \, dxdy = 0 \qquad (11f)$$

For reasons of symmetry one finds from (11b) $\lambda_2 = 0$, from (11c) $\lambda_3 = 0$, from (11e) $\lambda_5 = 0$. If we consider a region A which does not change if the x and y axis are interchanged, one furthermore finds $\lambda_4 = \lambda_6$. Let $\lambda_1 = 1$. We now have

$$\theta = 1 + \lambda_6 (x^2 + y^2)$$

with the additional conditions

$$\iint_A \theta x^2 \, dxdy = 0$$

$$\iint_A \theta y^2 \, dxdy = 0$$

or

$$\iint_A \theta (x^2 + y^2) \, dxdy = 0$$

If we consider a circular region of radius 1,* then this condition will give

$$\int_0^1 (1+\lambda_6 r^2)\, r^3\, dr = 0$$

$$\lambda_6 = -3/2$$

Thus

$$\theta = 1 - \frac{3}{2} r^2$$

With this value one obtains for the average measurement

$$M_Y = 2\pi k\rho_0 \int_0^1 (1 - \frac{3}{2} r^2)\, r dr = \frac{1}{4} \pi k\rho_0$$

The variance is

$$V_Y = 2\pi k\rho_0 \int_0^1 (1 - \frac{3}{2} r^2)^2\, r dr = \frac{1}{4} \pi k\rho_0$$

The number of points in the region is $N = \pi k\rho_0$

Thus the average of the measured quantity is $N/4$

and the variance is $N/4$

The relative error is

$$\epsilon = \frac{2\tau}{\sqrt{N}}$$

*Since the scale of the x and y coordinates should be unessential, it can always be chosen in such a manner, that at the boundaries $x = 1$, or $y = 1$, or $r = 1$ (whichever is convenient).

If one would have omitted the conditions that the result has to be sensitive against contributions of x^2 and y^2 in ρ then the relative error would have been 1/2 of this amount, viz

$$\epsilon = \frac{\tau}{\sqrt{N}}$$

For comparison the same computation may be carried out for a square region.

$$\theta = 1 + \lambda_6 (x^2 + y^2)$$

λ_6 is determined from the condition

$$\iint_A \theta x^2 \, dxdy = 0$$

Thus

$$\iint_A [x^2 + \lambda_6 (x^4 + x^2 y^2)] \, dxdy = 0$$

$$\lambda_6 = -\frac{15}{14}$$

Thus

$$\theta = 1 - \frac{15}{14} (x^2 + y^2)$$

$$M_Y = 4k\rho_0 \int_0^1 \int_0^1 \theta \, dxdy = 4k\rho_0 \int_0^1 \int_0^1 [1 - \frac{15}{14} (x^2 + y^2)] dxdy = \frac{8}{7} k\rho_0$$

The number of points in this region is $4k\rho_0 = N$

Thus

$$M_Y = \frac{2}{7} N$$

The variance is found to be

$$V_Y = 4k\rho_0 \int_0^1\int_0^1 [1 - \frac{15}{14}(x^2 + y^2)]^2 dxdy = \frac{8}{7} k\rho_0 = M_Y$$

Here the relative error is given by

$$\epsilon = \frac{\sqrt{3.5\tau}}{\sqrt{N}}$$

i. e. for the square the relative error is somewhat smaller than for a circle if the same number of points is used.

The limitations of this procedure are quite obvious. While in principle it is possible to exclude by constraints the influence of additional terms in a development for ρ each constraint may make the function θ more eneven and thus increase the variance. We shall take up this question later.

To show another variation of this technique let us try to determine a procedure for finding the first derivative of a given function in the x direction.

Given

$$\rho = \rho_0 (1 + a_1 x + a_2 y) \tag{12}$$

Assume again that a_1 and a_2 are small, so that $\rho \sim \rho_0$. Then an optimal expression θ has the form

$$\theta = \lambda_1 + \lambda_2 x + \lambda_3 y$$

subject to the conditions

$$\iint_A \theta x \, dxdy = 1$$

$$\iint_A \theta \, dxdy = 0$$

$$\iint_A \theta y \, dxdy = 0$$

The second condition gives immediately $\lambda_1 = 0$, the third one $\lambda_3 = 0$. Let the test area be a square of side $2a$. The first condition yields

$$\theta = \frac{3}{4a^4} x \tag{13}$$

One then finds

$$M_Y = \int_{-a}^{+a} \int_{-a}^{+a} \theta(x)\, \rho(x)\, dxdy = a_1 k \rho_0$$

Let $N' = k\rho_0$ be the expected number of points per unit of area (cf. (1)). Then

$$M_Y = a_1 N'$$

Hence

$$a_1 = \frac{M_Y}{N'} \tag{14}$$

The variance is

$$V_Y = \int_{-a}^{+a}\int_{-a}^{+a} \theta^2(x,y)\,\rho(x,y)\,dxdy = \frac{3}{4}\,\frac{k\rho_0}{a^4} = \frac{3}{4}\,\frac{N'}{a^4} \tag{15}$$

According to the remarks in Section 2 the half width of a region which contains a fraction $1 - \alpha$ of all measurements is $\tau\sqrt{V_Y}$. The uncertainty in the measured values of Y can be expressed as an uncertainty in the values of a_1. If the measurement Y deviates from M_Y by $\tau\sqrt{V_Y}$ then one obtains from (14) as deviation in a_1

$$\Delta a_1 = \frac{\tau\sqrt{V_Y}}{N'} = \frac{\tau\sqrt{3}}{2a^2\sqrt{N'}}$$

For $\alpha = 0.1$ (cf. Section 2) one finds $\tau = 1.6$ and

$$\Delta a_1 = \frac{2.76}{2a^2\sqrt{N'}} \tag{16}$$

For $a = 1$ the average number of points in the test area is $4N'$. Assume $4N' = 900$. Then

$$\Delta a_1 = 0.092$$

The significance of a_1 can be seen from Eq (12). The uncertainty in a_1 just found is rather large. Eq (16) shows how the size of the test area will influence the accuracy.

4. More general distribution of ρ

In previous considerations we assumed a certain form of the functions and then determined the accuracy with which the value at a given point could be obtained. The result was invariably that the region whould be taken as large as possible. A limit arises naturally by the fact that the analytic expressions for ρ used here will represent the actual ρ (x, y) in a limited region only. For functions ρ as they occur in practice the expressions

$$M_Y = \iint_A \theta(x, y)\, \rho(x, y)\, dxdy$$

are only approximations of the desired quantities ρ or grad ρ. To extend the techniques used so far to more and more complex analytical expressions for ρ would be rather useless, for with this process the variance will increase. Therefore we must accept the fact that the average values M_Y which we determine are only approximations of the desired quantities.

In the following discussions we shall use for the determination of ρ at a given point an approximation which is based on the assumption that in the area considered ρ is a linear function of x and y. Correspondingly $\theta = 1$. The value ρ in the center of the area is then approximately proportional to

$$M_Y = \int_{-a}^{+a} \int_{-a}^{+a} \rho(x, y)\, dxdy$$

where the origin of the xy - system lies at the center of the region considered.

On a given photograph there exist lines of constant ρ. Let us consider along one of these lines, say $\rho = \rho_0$, a great number of equidistant points. These points may be numbered as 1, 2 . . . j . . . their coordinates will be $(x_1, y_1) \ldots (x_j, y_j) \ldots$. We then obtain for the value M_Y for point j

$$M_{Y,j} = \int_{y_j - a}^{y_j + a} \int_{y_j - a}^{y_j + a} \rho(x, y)\, dxdy$$

Since the value of M_Y will vary from point to point (although all points j lie on a line ρ = constant), one may regard M_Y itself as a random function. For this random function there exists a probability distribution, which gives the probability density for M_Y to take on given values. For our measurements these values M_Y would be interpreted as proportional to ρ_0. Obviously this probability distribution and its variance is determined by the character of the object. Intuitively it is clear that for an object for which ρ varies only slowly, M_Y has a smaller variance than for an object with rather abrupt changes. In general the values of this variance must be estimated.

But it may be of some interest to see how this variance can be determined in principle.

Assume for this purpose that for the vicinity of the point (x_j, y_j) described above, ρ is actually of the form

$$\rho = \rho_0 [1 + a_{1,j}(x - x_j) + a_{2,j}(y - y_j) + a_{3,j}(x - x_j)^2 + a_{4,j}(y - y_j)^2]$$

The deviation of $M_{Y,j}$ from the value ρ_0 is easily computed as

$$M_{Y,j} - \rho_0 = \frac{4}{3}\rho_0(a_3 + a_4)a^4$$

The quantity $a_3 + a_4$ can be considered as a random variable which takes on the values $a_{3,j} + a_{4,j}$ for the points considered and possesses a variance $V_{(a_3+a_4)}$. We assume that the average of the random variable $a_3 + a_4$ is

$$M_{(a_3+a_4)} = \int_{-\infty}^{+\infty} (a_3 + a_4)\, p(a_3 + a_4)\, d(a_3 + a_4) = 0$$

The variance

$$V_{(a_3+a_4)} = \int_{-\infty}^{+\infty} (a_3 + a_4)^2\, p(a_3 + a_4)\, d(a_3 + a_4)$$

is considered as known.

The variance of M_Y about the correct value ρ_0 is then given by

$$V(M_y) = \int \rho_0^2 (a_3 + a_4)^2 \frac{16}{9} a^8 \; p(a_3 + a_4) \, d(a_3 + a_4)$$

hence

$$V(M_y) = \frac{16}{9} a^8 \rho_0^2 \, V_{(a_3 + a_4)}$$

The value of this calculation lies in the fact that it shows how $V(M_y)$ can be computed and how it depends upon "a" and other known quantities. Naturally for other assumed representations of ρ, a similar computation can be carried out.

$V(M_y)$ is the variance of the average measurement M_y from the light density ρ_0. For each value ρ_0 of the light density occurring in a given picture the variance $V(M_y)$ can be determined.

Since $4\rho_0 a^2$ is the number of points N in the region considered, the last result can also be written as

$$V(M_y) = \frac{N^2 a^4}{9} V_{(a_3 + a_4)}$$

The variance which occurs in the measurement of ρ has been shown in Section 3, Example a, to be N. According to the theorem quoted at the end of Section 2 the two variances must be added. Thus the total variance of the measurement of Y about ρ_0 is

$$N + \frac{N^2 a^4}{9} V_{(a_3 + a_4)}$$

The measure for the relative error is then

$$\epsilon = \frac{\tau \sqrt{N + \frac{N^2 a^4}{9} V_{(a_3 + a_4)}}}{N}$$

Since $N = 4\rho_0 a^2$, one has

$$\epsilon = \frac{\tau \sqrt{4\rho_0 a^2 + \frac{16}{9}\rho_0^2 V_{(a_3 + a_4)} a^8}}{4\rho_0 a^2} \tag{17}$$

To obtain best conditions this quantity must be minimized, i.e. one must determine the minimum of

$$\frac{4\rho_0}{a^2} + \frac{16}{9} V_{(a_3 + a_4)} a^4$$

Hence

$$a = \left(\frac{9}{8 V_{(a_3 + a_4)} \rho_0} \right)^{1/6} \tag{18}$$

This equation defines the best size of the region. One will notice that this result depends upon the value of ρ_0; therefore, "a" must be expected to vary with ρ_0. Among the values of "a" so obtained from one photograph one must choose a suitable value.

Since $V_{(a_3 + a_4)}$ is in general not known, this discussion is somewhat academic, but it shows how the size of the test region is connected to the character of the photograph.

No essential changes would occur if instead of (12) a different form of ρ would be assumed, e.g.

$$\rho = \rho_0(1 + a_1 x + a_2 y + a_3 |x| + a_4 |y|)$$

5. Scanning for Special Patterns

The examples up to now were concerned with the determination of ρ and of grad ρ. It may occur that certain special patterns of the light distribution are of importance and one might try to detect them. These patterns might be of such a nature that they would be averaged out if one simply determines ρ and tries to recognize them afterwards. An extreme example may be diffraction rings in a telescope due to the finiteness of the aperture. They may be rather completely hidden in the background light. More practical examples are sharp lines like streets or rivers which are so narrow that they become rather indistinct. Also the question may arise whether along a certain line the density jumps, or whether there is a gradient. Such cases are covered by our previous investigations. If the function ρ_1

describes the pattern which is to be recognized, and ρ_2 and ρ_3 are functions which should be suppressed, and if ρ is nearly constant, then the function $\theta(x,y)$ has the form

$$\theta = \lambda_1 \rho_1 + \lambda_2 \rho_2 + \lambda_3 \rho_3$$

and the λ's are to be determined from the conditions that the influence of ρ_2 and ρ_3 must be suppressed.

Assume e.g. that ρ_1 is the distribution due to the Newton ring, and that $\rho_2 = 1$ is the background light.

Let

$$\rho = \rho_0 (1 + a_1 \rho_1)$$

then one has $\theta = \rho_1 + \lambda \rho_2$

and from $\iint_A \theta \rho_2 \, dxdy = 0$

one obtains, setting $\rho_2 = 1$

$$\lambda = -\frac{1}{A} \iint_A \rho_1 \, dxdy$$

i.e. θ is equal to ρ_1 minus a constant which makes the average of θ over A zero.

The average of the measurements is then

$$M_Y = k \rho_0 a_1 \iint_A (\rho_1 + \lambda \rho_2) \rho_1 \, dxdy \tag{19}$$

The variance of the measurement is

$$V_Y = k\rho_0 \iint_A (\rho_1 + \lambda \rho_2)^2 \, dxdy \qquad (20)$$

To get a feel for the orders of magnitude let us assume that the area considered is A, and that the function ρ_1 is 1 in the area γ_0 A and is zero in the rest of the region, where γ_0 is a constant between 0 and 1 (See Fig 2.) Further, for simplicity, let a_1 be small. Then

$$\iint_A \rho_1 \, dxdy = \gamma_0 A$$

and

$$\lambda = - \gamma_0$$

$$\theta = \rho_1 - \gamma_0 \rho_2$$

The average measurement will give

$$M_Y = k\rho_0 a_1 \iint_A (\rho_1 - \gamma_0 \rho_2)\, \rho_1 \, dxdy$$

$$= k\rho_0 a_1 \, A \gamma_0 (1 - \gamma_0) = N a_1 \gamma_0 (1-\gamma_0)$$

Let $N' = N\gamma_0$ be the number of points in the area where $\rho_1 \neq 0$.

Then $M_Y = a_1 (1 - \gamma_0) N'$

The variance is

$$V_Y = k\rho_0\, \gamma_0\,(1-\gamma_0) = N\gamma_0\,(1-\gamma_0) = N'(1-\gamma_0)$$

We proceed by analogy to the treatment of the density gradient. The half width of a region in which a fraction $(1-\alpha)$ of all measurements Y are contained is given by $\tau\sqrt{V_Y}$.

One may ask for which value of a_1 the half width equals the expected value M_Y; one finds

$$a_1 = \frac{\tau}{\sqrt{1-\gamma_0}\,\sqrt{N'}} \tag{21}$$

The aim should be to make this value small. One limited possibility is the choice of γ_0. However, if we keep N' fixed and choose γ_0 small, then we increase the total test area, while the area belonging to the pattern is the same. In this process we would encounter additional errors due to the nonuniformity of the background field. A value of $\gamma_0 = 1/2$ appears reasonable. Using $\tau = 1.6$ (corresponding to $\alpha = 0.1$) one obtains

$$a_1 = \frac{1.6}{\sqrt{1/2}\,\sqrt{N'}}$$

For $a_1 = 0\ 1$, i.e. for a pattern which is 10% darker than the background, the pattern must contain $N' = \frac{2(1.6)^2}{a_1^2} = 512$ points to meet this condition.

This approach applies to patterns of any kind. One might, e. g. try to recognize a straight line of points arranged over a certain width. In this case one would consider a test area of about twice the expected width extending in the direction of the expected straight line. (This description shows that, for the purpose, the testing will be done with a rather narrow rectangular test area, which must eventually be rotated if one tries to scan for straight lines of different direction.) Eq (21) applies immediately and allows us to determine the limit where a pattern can still be recognized.

One may ask, what will happen if one tests for a pattern which differs from the pattern which occurs in the picture. Let us assume that the line where the darkness occurs has actually one half the width of what we expect it to have and that its density is twice what we expect it to have, i. e. that the value a_1 is now twice the value of the previous case. Then since the number of grains in A and in the pattern is the same as before, one obtains from the calculations the same value a_1 as before. In reality a_1 is twice this calculated value. The variance has the same value as before. Furthermore, if the proper width had been used, then N' would be one-half of the original value and the true value of a_1 now is $\sqrt{2}$ times the calculated value of a_1. This shows that the scanning should be adapted to the pattern as well as possible.

A step in ρ in the x direction may be described by

$$\rho_1 = -1 \text{ for } x < 0$$

$$\rho_1 = +1 \text{ for } x > 0$$

To scan for such a step one must choose a function θ which will not register a constant background density, or a density gradient, or a step in the y direction.

Let

$$\rho_2 = 1$$

$$\rho_3 = x$$

$$\rho_4 = y$$

$$\rho_5 = -1 \text{ for } y < 0$$

$$\rho_5 = +1 \text{ for } y > 0$$

Then one has

$$\theta = \rho_1 + \lambda_2\rho_2 + \lambda_3\rho_3 + \lambda_4\rho_4 + \lambda_5\rho_5$$

with the additional conditions

$$\iint_A \theta \, \rho_1 \, dxdy \neq 0 \qquad (22a)$$

$$\iint_A \theta\, \rho_i \, dxdy = 0\,, \quad i = 2, 3 \ldots 5 \tag{22b}$$

These are 4 conditions determining $\lambda_2 \ldots \lambda_5$. All λ's except λ_3 turn out to be zero. This is seen if one assumes θ to have the form

$$\theta = \rho_1 + \lambda\, \rho_3 \tag{23}$$

Since ρ_1 and ρ_3 are odd functions of x while ρ_2, ρ_4 and ρ_5 are even functions of x, it is immediately obvious that condition (22b), is fulfilled by (23).

For an interval $-a < x < a$, one obtains from (22c)

$$\int_{-a}^{+a} \theta\, x\, dx = 0$$

and hence

$$\lambda = -\frac{3}{2} \quad \frac{1}{a}$$

In determining the mean value we can omit all contributions to ρ which is due to $\rho_2 \ldots \rho_5$. Let

$$\rho = \rho_0 \left(1 + a_1\, \rho_1 \right)$$

The average measurement is

$$\begin{aligned} M_Y &= k a_1 \rho_0 \iint_{-a\,-b}^{+a\,b} \rho_1 \left(\rho_1 - \frac{3}{2} \, \frac{x}{a} \right) dxdy \\ &= k a_1 \rho_0 \; ab = \frac{a_1 N}{4} \end{aligned}$$

In determining the variance it is assumed that ρ is nearly constant; actually this assumption is not needed with the present choice of ρ_5. One obtains,

$$V_Y = 2k\rho_0 \int_0^a \int_{-b}^{b} \left(1 - \frac{3}{2}\frac{x}{a}\right)^2 dxdy$$

$$= \frac{N}{4}$$

The discussion carried out in conjunction with (21) applies again. In order to obtain the same limit where a pattern can still be recognized, the number N in the present case must be $4(1 - \gamma_0)$ times as large as before.

Quite obviously, if the step in the photograph occurs at a position different from $x = 0$ and we test for a step at $x = 0$, we shall obtain a result which indicates the presence of a step. The same applies if the step does not extend parallet to the y-axis, but is inclined to it. In principle, it might be feasible to add conditions which would guard against misinterpretations of this kind, but because of the large number of grains required, the practical value of such a method is doubtful. One would probably make the test indicated here for different locations of the jump and for different directions of it and then place the jump in the position where the indication is maximum.

6. Recognition of Darker Spots Against a Uniform Background

The previous section dealt with the problem of measuring certain defined quantities and of determining confidence limits for these measurements. In this section we treat a case where we know that the light density can assume two values, that of the background and that of the signal which raises the light density slightly above the background. We want to discriminate between areas of this background light density and of the higher light density.

To be specific, let us assume that the plate is divided into a rather large number of areas of equal size (e. g. squares) and that these squares have been exposed to the background light or to the background light plus the light of the signal. In a fraction β of all squares we have the background light and corresponding to it a value N_1 as expected number of grains; in the remaining fraction $(1 - \beta)$ we have background plus signal and correspondingly an expected number of grains N_2. For convenience of discussion, if an area has an expected number of grains equal to N_1, we shall say it "belongs to N_1". Let $N_1 < N_2$. The regions corresponding to the signal (N_2) are to be found. For this purpose we count the number of grains in each area and consider the areas below a certain cut off count m as belonging to N_1, all areas above the cut off count as belonging to N_2.

m should be chosen in such a manner that the probable number of wrong decisions is minimized. (This formulation is not the only possible one, one might e. g. count an error committed by missing a dark spot more heavily than an error where a light spot is considered as dark.)

The probability of finding n grains in an area which belongs to an expected value N_1 is given by a Poisson distribution. Therefore, if m is the cutoff count, the probability of misjudging areas belonging to N_1 as areas belonging to N_2 is given by

$$\beta\, e^{-N_1} \sum_{m}^{\infty} \frac{N_1^n}{n!}$$

The probability of misjudging areas belonging to N_2 as areas belonging to N_1 is

$$(1-\beta)\, e^{-N_2} \sum_{-\infty}^{m-1} \frac{N_2^n}{n!}$$

Thus the probability of making an error is

$$p_e = \beta\, e^{-N_1} \sum_{m}^{\infty} \frac{N_1^n}{n!} + (1-\beta)\, e^{-N_2} \sum_{-\infty}^{m-1} \frac{N_2^n}{n!} \tag{24}$$

To find a condition for m, we increase and decease the cutoff count m by one; in either case the error probability must rise. In the first case the error probability changes by

$$-\beta\, e^{-N_1} \frac{N_1^m}{m!} + (1-\beta)\, e^{-N_2} \frac{N_2^m}{m!} \tag{25}$$

In the second case it changes by

$$- \left\{ -\beta\, e^{-N_1} \frac{N_1^{m-1}}{(m-1)!} + (1-\beta)\, e^{-N_2} \frac{N_2^{m-1}}{(m-1)!} \right\} \tag{25b}$$

If m corresponds to a minimum probability of making an error both expressions must be positive. The second expression originates from the first one by a sign change and by replacing m by m-1. If one replaces m by x and considers (25a) as a function of the continuously varying variable x, then this function must have a zero between m-1 and m. From this condition the cutoff count can be determined. From the equation

$$-\beta\, e^{-N_1} \frac{N_1^{x}}{x!} + (1-\beta)\, e^{-N_2} \frac{N_2^{x}}{x!} = 0$$

one obtains

$$x = x' + x''$$

where

$$x' = \frac{\ln \frac{1-\beta}{\beta}^{\,*}}{-\ln \frac{N_2}{N_1}} \quad , \quad x'' = \frac{N_2 - N_1}{\ln \frac{N_2}{N_1}}$$

* The symbol ln represents "the natural logarithm of"

In practice it may be possible to change the size of the test areas while the ratio N_2/N_1 is fixed. Of particular interest is the case where N_2/N_1 is close to one. Let

$$\frac{N_2}{N_1} = \gamma_1$$

Then

$$x' = \frac{\ln \frac{1-\beta}{\beta}}{-\ln \gamma_1}{}^{*}$$

The expression x' does not depend explicitly upon N_1 or N_2, only upon their ratio. It gives a constant shift in the cutoff count which is independent of the size of the test areas (which is characterized by N_1 or N_2). The quantity x'' is best described in an intuitive manner by the expression

$$\delta = \frac{x'' - N_1}{N_2 - N_1}$$

i.e. by the quotient of $x'' - N_1$ and $N_2 - N_1$. One finds

$$\delta = \frac{1}{\ln \gamma_1} - \frac{1}{\gamma_1 - 1}$$

This quantity is shown in Fig. 3; it is rather close to $\frac{1}{2}$.

Of prime interest are actual values for the probability of making errors. To compute these the Poisson distributions in (24) are replaced

*The symbol ln represents "the natural logarithm of" by the corresponding normal distributions, and the sums by integrals, (the admissibility of this approximation is shown in Appendix II). One then finds

$$P_e = \beta \frac{1}{\sqrt{2\pi N_1}} \int_{m-\frac{1}{2}}^{\infty} e^{\frac{-(x+\frac{1}{2}-N_1)^2}{2N_1}} dx + (1-\beta) \frac{1}{\sqrt{2\pi N_2}} \int_{-\infty}^{m-\frac{1}{2}} e^{\frac{-(x+\frac{1}{2}-N_2)^2}{2N_2}} dx$$

Introducing in the first integral

$$u = \frac{x+\frac{1}{2}-N_1}{\sqrt{N_1}}$$

and in the second integral

$$u = \frac{x+\frac{1}{2}-N_2}{\sqrt{N_2}}$$

one obtains

$$P_e = \frac{1}{\sqrt{2\pi}} \quad \beta \int_{u_1}^{\infty} e^{-\frac{u^2}{2}} du + (1-\beta) \int_{-\infty}^{u_2} e^{-\frac{u^2}{2}} du$$

$$\text{where } u_1 = x' N^{-\frac{1}{2}} + \delta(\gamma_1 - 1) N_1^{\frac{1}{2}}$$

$$u_2 = x' \gamma_1^{-\frac{1}{2}} N_1^{-\frac{1}{2}} - (1-\delta)(\gamma-1)\gamma_1^{-\frac{1}{2}} N_1^{\frac{1}{2}}$$

The values for these integrals in terms of the upper limit can be found in tables. Some examples are given in Figs. 4 through 12. In the examples $\gamma_1 = \frac{N_2}{N_1} = 1.1,\ 1.5,$ and 2.0. For β the values 0.1, 0.5, and 0.9 have been chosen. The value of N_i varies from 50 to 1000. The probability of making an error should be compared with $(1-\beta)$ which indicates in how many of the cases a signal is present. As β is reduced this ratio becomes rather high although the probability of making an error is reduced.

7. A Comparison of Two Techniques of Assigning a Light Density to a Single Area of Lesser Intensity than the Surroundings

Assume now that we look at a given area, a portion of the photographic plate, which we shall designate as the "test area". The number of grains which we count is n. We want to decide whether this test area corresponds to an expected number N_1 or to some alternative expected number N_2, $N_2 > N_1$ (without specifying what N_2 is).

Two different strategies for making this decision will be compared: We choose a number r, $r>1$. In Strategy 1, we count the number of grains in the test area. If it is below rN_1 we consider this area as belonging to N_1. In Strategy 2, the test area is subdivided into smaller areas (for the sake of argument, into nine subareas, let us say). Let $N_1' = N_1/9$. We choose another number r', $r'>1$. Then we determine for each of the smaller areas whether the count is smaller than $r'N_1'$. To the original test area we then assign the count N, if more than half of the subareas have a count smaller than $r'N'$.

Several questions immediately arise:

(1) How shall we choose r and r' ?

(2) Which strategy is superior ?

(3) How much better quantitatively is the superior method ?

Intuitively one might consider Strategy 1 (using the whole area) as superior; however Strategy 2 (by subareas) may recommend itself for practical reasons. The object of this section is to determine quantitatively r, r' and the probabilities of error in either case.

As stated in Section 6 (cf p. 35), there are two possibilities for making a wrong decision:

(1) misjudging the test area as not belonging to N_1 when it actually does belong to N_1 (This is called "Type I error");

(2) misjudging the test area as belonging to N_1 when it actually belongs to some other number N_2, N_2 different from N_1 (This is called "Type II error"). In this section N_2 is greater than N_1.

For a number N_1 (or equivalently, N_1') and a fixed value of r (or r') one can calculate the probability of making a wrong decision of Type 1 for each of the strategies. The probability of making a wrong decision of Type 2 is then a function of N_2 for each of the strategies.

Intuitively one can see that the closer N_2 is to N_1 the greater is the probability of making a wrong decision.

Consider Strategy 1 (whole area method). We assume that the energy density of the incoming light is constant over the test area. To fix the argument let us temporarily assume that the test area has an expected value of N_1. Under Strategy 1 we will decide that the test area does not belong to N_1 if the number of grains counted is greater than rN_1. Let P be the probability that we misjudge the test area (with expected value N_1) as not belonging to N_1, using Strategy 1.

Then

$$P = e^{-N_1} \sum_{[rN_1+1]}^{\infty} \frac{(N_1)^n}{n!}$$

where the brackets indicate "largest integer in".

Then

$$P \approx \frac{1}{\sqrt{2\pi N_1}} \int_{rN_1}^{\infty} e^{-\frac{(x-N_1)^2}{2N_1}} dx \quad \text{for } N_1 \geq 5$$

Let $u = \dfrac{x-N_1}{\sqrt{N_1}}$

Then P is approximately given by

$$P = \frac{1}{\sqrt{2\pi}} \int_{(r-1)\sqrt{N_1}}^{\infty} e^{-u^2/2} du = \frac{1}{\sqrt{2\pi}} \int_{3(r-1)\sqrt{N_1'}}^{\infty} e^{-u^2/2} du$$

Column 4 of Table 1 gives values for P corresponding to $(r-1)\sqrt{N_1'} = 0(.1)1.2$.

Now consider Strategy 2 (subarea method). If the expected number of grains in the test area is N_1, then the expected number of grains in each subarea is N_1' where $N_1' = N_1/9$, and the probability of finding n grains in any subarea is given by a Poisson distribution with $\lambda = N_1'$. Under Strategy 2 we will decide that a subarea does not belong to N_1' if the number of grains counted is greater than $r'N_1'$. Let p_s be the probability that we misjudge a subarea (with expected value N_1') as not belonging to N_1'. The subscript s denotes the subarea method.

Then

$$p_s = e^{-N_1'} \sum_{[r'N_1'+1]}^{\infty} \frac{(N_1')^n}{n!}$$

$$\approx \frac{1}{\sqrt{2\pi N_1'}} \int_{r'N_1'}^{\infty} e^{-\frac{(x-N_1')^2}{2N_1'}} dx \quad \text{for } N_1' \geq 5$$

Let $u = \dfrac{x-N}{\sqrt{N_1'}}$

Then

$$p_s = \frac{1}{\sqrt{2\pi}} \int_{(r'-1)\sqrt{N_1'}}^{\infty} e^{-u^2/2} du$$

In the case $r' = r$, Column 2 of Table I gives values of p_s for

$(r-1)\sqrt{N'} = 0(.1)1.2$.

Further, we will decide that the original test area does not belong to N_1 if more than half of the subareas have a count greater than $r'N'_1$. Let P_s be the probability (using the subarea method) that we misjudge the test area (with expected value N_1) as not belonging to N_1. Since the number of grains counted in each subarea is independent of that for any other subarea, this probability is a sum of binomial terms, viz:

$$P_s = p_s^9 + \binom{9}{1} p_s^8 (1-p_s) + \binom{9}{2} p_s^7 (1-p_s)^2 + \binom{9}{3} p_s^6 (1-p_s)^3 + \binom{9}{4} p_s^5 (1-p_s)^4$$

where p_s^9 is the probability that all nine subareas with expected value N'_1 are misjudged as not belonging to N'_1;

$\binom{9}{1} p_s^8 (1-p_s)$ is the probability that eight subareas are misjudged as not belonging to N'_1 (and one subarea is correctly judged as belonging to N'_1);

....

$\binom{9}{4} p_s^5 (1-p_s)^4$ is the probability that five subareas are misjudged as belonging to N'_1 (and the other four subareas are correctly judged as belonging to N'_1).

There are several possible ways to obtain P_s once p_s is known. One is to interpolate between values given in a table such as the

Tables of the Cumulative Binomial Probability Distribution (Harvard University, Cambridge University Press, 1955). Another possibility is to interpolate in a table such as Tables of the Incomplete Beta-function Ratio (edited by Karl Pearson, Biometrika Office, University College, London 1934) since it can be shown (see Kendall [1], p. 120) that

$$P_s = I_{p_s}(5, 5) = 1 - I_{1-p_s}(5, 5)$$

where $I_x(p, q)$ is the incomplete Beta-function ratio. However, here it was more convenient to calculate P_s directly from p_s with the use of an E-101 Burroughs electronic computer. The results are given in Table I. For example, when $r' = r$, Column 3 of Table I gives values of P_s corresponding to $(r-1)\sqrt{N'} = 0\ (0.1)\ 1.2$.

The values of P and p_s can be obtained by using a table such as Tables of Probability Functions, Volume II (Mathematical Tables Projects, National Bureau of Standards, 1942).

For the case that $r' = r$, the last column of Table I compares the Type I error probability of Strategy 1 (whole area method) with that of Strategy 2 (subarea method) in the form of a ratio, P/P_s. Strategy 1 gives lower error probabilities for all values listed; the difference is more pronounced if the Type I error probability is low (50% error means no judgement at all). Let $r = 1.05$. This would correspond to a case where we want to discriminate between two light densities which differ by 10% and where we lay the cut off grain count into the middle. Let $N' = 100$, which corresponds to 900 points in the larger square; then $(r-1)\sqrt{N'} = .5$ and the probability of making a Type I error is 0.1096 with Strategy 2 versus .066= with Strategy 1.

Thus far we have discussed only probabilities of wrong decisions of Type I for the two strategies (see page 41). Consider now the wrong decisions of Type II for the two strategies, i.e. misjudging the test area with expected value N_2, $N_2 > N_1$, as belonging to N_1.

For Strategy 1 (whole area method), we now consider the test area to have an expected value N_2, where $N_2 = \gamma_1 N_1$ $\gamma_1 > 1$. Under Strategy 1 we will decide that a subarea belongs to N_1 if the number of grains counted is less than rN_1. Let Q be the probability that we misjudge the test area (with expected value N_2) as belonging to N_1.

Then

$$Q = e^{-N_2} \sum_{0}^{[rN_1]} \frac{(N_2)^n}{n!}$$

$$\approx \frac{1}{\sqrt{2\pi N_2}} \int_{-\infty}^{rN_1} e^{-\frac{(x-N_2)^2}{2N_2}} dx \quad \text{for } N_2 \geq 5$$

Let $u = \frac{x - N_2}{\sqrt{N_2}}$

Then $$Q \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{(r-\gamma_1)\sqrt{N_1}}{\sqrt{\gamma_1}}} e^{-u^2/2}\, du \qquad (5)$$

where $N_2 = \gamma_1 N_1$

For Strategy 2, a subarea has an expected number of grains N_2', where $N_2' = \gamma_1 N_1/9$, and the probability of finding n grains in any subarea is a Poisson distribution with $\lambda = N_2'$. Under Strategy 2 we will decide that a subarea belongs to N_1' if the number of grains counted is less than $r'N_1'$. Let q_s be the probability that we misjudge a subarea with expected value N_2' as belonging to N_1'.

Then

$$q_s = e^{-N_2'} \sum_{0}^{[r'N_1']} \frac{(N_2')^n}{n!}$$

or $$q_s \approx \frac{1}{\sqrt{2\pi N_2'}} \int_{-\infty}^{r'N_1'} e^{-\frac{(x-N_2')^2}{2N_2'}}\, dx \qquad \text{for } N' \geq 5$$

Let $$u = \frac{x-N_2'}{\sqrt{N_2'}}$$

Then

$$q_s \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_1} e^{-\frac{u^2}{2}}\, du = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_2} e^{-u^2/2}\, du \qquad (6)$$

$$\text{where } u_1 = \frac{r_1'N_1' - N_2'}{\sqrt{N_2'}}, \quad u_2 = \frac{(r'-\gamma_1)\sqrt{N_1}}{3\sqrt{\gamma_1}}$$

Furthermore, we will decide that the original test area belongs to N_1, if more than half of the subareas have a count less than $r'N_1'$. Let

Q_s be the probability (using the subarea method) that we misjudge the test area (with expected value N_2) as belonging to N_1.

Then, this probability is again a sum of binomial terms:

$$Q_s = q_s^9 + \binom{9}{1} q_s^8 (1-q_s) + \binom{9}{2} q_s^7 (1-q_s)^2 + \binom{9}{3} q_s^6 (1-q_s)^3$$

$$+\binom{9}{4} q_s^5 (1-q_s)^4$$

where q_s^9 is the probability that all nine subareas with expected value N_2' are misjudged as belonging to N_1';

$\binom{9}{1} q_s^8 (1-q_s)$ is the probability that eight subareas with expected value N_2' are misjudged as belonging to N_1' (and one subarea is correctly judged as not belonging to N_1');

....

$\binom{9}{4} q_s^5(1-q_s)^4$ is the probability that five subareas are misjudged as belonging to N_1' (and four subareas are correctly judged as not belonging to N_1').

One again has the relationship:

$$Q_s = I_{q_s}(5,5) = 1 - I_{q_s}(5,5)$$

where $I_x(p,q)$ is the incomplete Beta-function ratio.

We now have the necessary formulas to enable us to compare the effectiveness of the two methods for deciding whether the light density of the test area is N_1. To make the discussion more explicit let us fix the value of γ_1 at $\gamma_1 = 1.1$. We set the probabilities of errors of Type I, P and P_s, at some common small positive value (less than one). The values .10, .05, and .01 are used here. Then we solve for the corresponding values $(r-1)\sqrt{N_1'}$ and $(r'-1)\sqrt{N_1'}$. (For a given value of N_1, we can then find the appropriate r and r'. Values of r and r' are given in Tables II(a) and II(b), respectively, for $P = P_s = .10, .05, .01$, with values of N_1 ranging from 5 to 1000 for r'.

With these values of r and r', and a value of γ_1, we then calculate the probabilities of making Type II errors, Q and Q_s, for the two methods (using the above formulas). Values of Q and Q_s are listed in Table III for $P = P_s = .10, .05, .01$, with values of γ_1 ranging from 1.1 to 2.0 and values of N_1 ranging from 50 to 1000.

Note that Q is always less than Q_s for these values of α. This evidence tells us that Strategy 1 using the whole area is the better method.

As an illustration of a use of Tables IIa and b and Table III, consider $\gamma_1 = 1.1$. This corresponds to the case where the background light intensity is 10% darker. If we set $r = 1.05$ (i.e., we lay the cutoff grain count for the whole-area method in the middle) and $P = .05$ then the table for r yields $N_1 = 1000$. And the corresponding value of r' (the cutoff grain count for the subarea method) would be 1.06 for $P_s = .05$. Thus, if the probability of Type I error is set at .05 for both methods, the probability of making Type II error is .1840 for Strategy 2 as compared with .0740 for Strategy 1.

One can proceed further, using the formulas for P and Q (for the whole area method), for example, in the following manner:

> Set P and Q at pre-assigned small values. Then one obtains a relationship between γ_1 and N_1.

For the case $P = Q = .05$ one has:

$$(r - 1)\sqrt{N_1} = 1.6448\overset{+}{5}$$

and

$$\frac{r - \gamma_1}{\sqrt{\gamma_1}} \sqrt{N_1} = -1.64485$$

Eliminating r from the above two equations and solving for γ_1 one obtains

$$\gamma_1 = \left(1 + \frac{1.6448\overset{+}{5}}{\sqrt{N_1}}\right)^2$$

Figure 13 gives the graph of N_1 vs. γ_1 using the above equation. N_1 as plotted corresponds to the minimum expected number of grains for the test area which is required such that neither of the probabilities of Type I and Type II errors exceeds .05 (for the whole area method). For example, consider $\gamma_1 = 2.0$. This corresponds to the situation where the surroundings are twice as dark as the test area. The graph yields 16 as the minimum required value for the expected number of grains in the test area. The values of γ_1 corresponding to various values of N_1 from 5 to 1000 is given in the table below.

N_1	γ_1
5	3.012
10	2.311
20	1.871
30	1.691
40	1.588
50	1.519
100	1.356
200	1.246
500	1.153
1000	1.107

Another possible application of the above formula for P is the determination of a 100(1-P)% confidence lower limit for N_1. This can be obtained by a plot of rN_1 vs N_1 corresponding to a pre-assigned value of P. Then for a given value of n, the observed number of grains (taken on the rN_1 axis),one obtains L_{N_1}, a lower 100(1-P)% confidence limit for N_1 (taken on the N_1 axis.) Figure 14 gives the graph of L_{N_1} vs n for P = .05 (95% confidence). To illustrate the use of such a graph suppose that 10 points are counted in the test area, i.e. $n = 10$. For $P = .05$, the corresponding value for L_{N_1} is 6. Hence a lower 95% confidence limit for the expected number of grains in the test area is 6. In layman's language, this means that if we count 10 points in the test area we can expect (with a 95% chance of being right in the long run) the average number of grains in the test area to be greater than or equal to 6.

If so desired, one may obtain upper $100(1-P)\%$ confidence limits for N_1 by considering $\gamma_1 < 1$, and using a technique analogous to that described above in this section.

8. Assigning a Light Density to a Single Area when Its Surroundings Contain Both Lesser and Greater Intensities

As in Section 7, let the number of grains which we count in the test area be n. We wish to decide whther this test area corresponds to an expected number N_1 or to an alternative number N_2 (without specifying N_2). However, now the only restriction on N_2 is that it is different from N_1.

The two strategies stated in Section 7 with suitable modifications may again be compared. Choose two numbers r_1, r_2 with $r_1 < 1, r_2 > 1$. In Strategy 1 (using the whole area) we count the number of grains in the test area. If it is greater than $r_1 N_1$ and less than $r_2 N_1$ we consider this area as belonging to N_1. Otherwise, we decide that the test area does not belong to N_1. For Strategy 2, suppose the test area is subdivided into nine subareas again. Choose two numbers r_1', r_2'

with $r_1' < 1$, $r_2' > 1$. Then we determine for each of the subareas whether the count is between $r_1' N_1/9$ and $r_2' N_1/9$. To the original test area we then assign the count N_1 if more than half of the subareas have a count between $r_1' N_1/9$ and $r_2' N_1/9$. One now has four numbers to choose: r_1, r_2, r_1', r_2'.

As in Section 7, it will be seen that Strategy 1 (using the whole area) is superior to Strategy 2. We follow the same development as given in Section 7 and obtain formulas for the Type I errors P, P_s and the Type II errors Q, Q_s.

Thus if P is the probability, using Strategy 1, that we misjudge the test area (with expected value N_1) as not belonging to N_1, then

$$P = e^{-N_1}\left\{\sum_{n=0}^{[r_1N_1]} \frac{(N_1)^n}{n!} + \sum_{n=[r_2N_1+1]}^{\infty} \frac{(N_1)^n}{n!}\right\}$$

$$\text{or} \quad P \approx \frac{1}{\sqrt{2\pi N_1}} \int_{-\infty}^{r_1 N_1} e^{-\frac{(x-N_1)^2}{2N_1}} dx + \frac{1}{\sqrt{2\pi N_1}} \int_{r_2N_1}^{\infty} e^{-\frac{(x-N_1)^2}{2N_1}} dx$$

$$\approx 1 - \frac{1}{\sqrt{2\pi}} \int_{(r_1-1)\sqrt{N_1}}^{(r_2-1)\sqrt{N_1}} e^{-\frac{u^2}{2}} du \qquad \text{for } N_1 \geq 5$$

It can be shown that good choices for pairs of values of r_1, r_2 are ones which yield symmetrical limits in the above integration, i.e. such that $(r_1-1) = -(r_2-1)$. We will consider only symmetrical limits here

Thus $$P \approx 1 - \frac{2}{\sqrt{2\pi}} \int_0^{(r_2-1)\sqrt{N_1}} e^{-\frac{u^2}{2}} du$$

Table IV gives values of r_1 and r_2 (for various values of N_1) such that $P = .10, .05, .01$.

Now if Q is the probability, using Strategy 1, that we misjudge the test area (with expected value N_2) as belonging to N_1, then

$$Q = e^{-N_2} \sum_{[r_1 N_1 + 1]}^{[r_2 N_1]} \frac{(N_2)^n}{n!}$$

or $$Q \approx \frac{1}{\sqrt{2\pi N_2}} \int_{r_1 N_1}^{r_2 N_1} e^{-\frac{(x-N_2)^2}{2N_2}} dx = \frac{1}{\sqrt{2\pi}} \int_{\frac{(r_1-\gamma_1)\sqrt{N_1}}{\sqrt{\gamma_1}}}^{\frac{(r_2-\gamma_1)\sqrt{N_1}}{\sqrt{\gamma_1}}} e^{-\frac{u^2}{2}} du$$

where $N_2 = \gamma_1 N_1$, $\gamma_1 > 0$,* and $N_2 \geq 5$

Substituting for r_1 in terms of r_2 [using the symmetrical relation $(r_1-1) = -(r_2-1)$] we obtain

$$Q \approx \frac{1}{\sqrt{2\pi}} \int_{\frac{-(r_2-1)\sqrt{N_1}}{\sqrt{\gamma_1}} - \frac{(\gamma_1-1)\sqrt{N_1}}{\sqrt{\gamma_1}}}^{\frac{(r_2-1)\sqrt{N_1}}{\sqrt{\gamma_1}} - \frac{(\gamma_1-1)\sqrt{N_1}}{\sqrt{\gamma_1}}} e^{-\frac{u^2}{2}} du \quad \text{for } \gamma_1 N_1 \geq 5$$

*Note that γ_1 may be less than or greater than one, since N_2 may be greater than or less than N_1 for this case.

As before, P_s is the probability, using Strategy 2, that we misjudge the test area (with expected value N_1) as not belonging to N_1. By analogy to the argument used in the previous section, we obtain the following formula for this case:

$$P_s = I_{p_s}(5, 5)$$

where $I_x(p, q)$ is the incomplete Beta-function ratio

$$\text{and} \quad p_s \approx 1 - \frac{2}{\sqrt{2\pi}} \int_0^{(r_2' - 1)\sqrt{N_1'}} e^{-\frac{u^2}{2}} du \qquad \text{for } N_1' \geq 5$$

As before, p_s is the probability that we misjudge a subarea (with expected value N_1') as not belonging to N_1'.

Finally, if Q_s is the probability, using Strategy 2, that we misjudge the test area (with expected value N_2) as belonging to N_1, then

$$Q_s = I_{q_s}(5, 5)$$

where $I_x(5, 5)$ is the incomplete Beta function ratio

$$\text{and } q_s \approx \frac{1}{\sqrt{2\pi}} \int_{\frac{(r_1' - \gamma_1)\sqrt{N_1'}}{\sqrt{\gamma_1}}}^{\frac{(r_2' - \gamma_1)\sqrt{N_1'}}{\gamma_1}} e^{-\frac{u^2}{2}} du = \frac{1}{\sqrt{2\pi}} \int_{-\frac{(r_2' + \gamma_1)}{\sqrt{\gamma_1}}\sqrt{N_1'}}^{\frac{(r_2' - \gamma_1)}{3\sqrt{\gamma_1}}\sqrt{N_1}} e^{-\frac{u^2}{2}} du$$

with $N_2 = \gamma_1 N_1$ and $N_2 \geq 5$

As before, q_s is the probability that we misjudge a subarea (with expected value N_2') as belonging to N_1'.

Proceeding as we did in Section 7, we set P and P_s (the Type I errors for the two methods) at some common value. Again the values .10, .05, and .01 are used here. Then we obtain the corresponding values of $(r_2-1)\sqrt{N_1}$ and $(r_2'-1)\sqrt{N_1'}$ using the formulas for P and P_s above. For a given value of N_1, we can then find the appropriate r_2 and r_2'. Furthermore, we can obtain the appropriate r_1 and r_1' using the symmetry relations.

Values of r_1, r_2 are given in Table IV(a) and values of r_1', r_2' are given in Table IV(b) for $P=P_s=.10$, .05, .01 with values of N_1 ranging from 5 to 1000 for (r_1, r_2) and from 50 to 1000 for (r_1', r_2').

For given values of r and r', and a value of γ_1, we can calculate Q and Q_s (the probabilities of Type II errors for the two methods). These are given in Table V for $P=P_s=.10$, .05, .01 with values of γ_1 ranging from 0.5 to 2.0 and values of N_1 ranging from 50 to 1000.

As in the previous section, an examination of Table V reveals that Q is always less than Q_s for the values considered. This evidence tells us that Strategy 1 (using the whole area) is, as expected, the better method. For example, consider $\gamma_1 = 0.9$. This corresponds to the case where the test area is 10% lighter than the background light intensity. If we set $r_1 = .94$, $r_2 = 1.06$ and $P=.05$ then $N_1 = 1000$. The corresponding values of r_1', r_2' (for $P_s=.05$) are $r_1' = .89$ and $r_2' = 1.11$. Thus, if the probability of Type I error is set at .05 for both methods, the probability of making Type II error is .10 for the whole area method as contrasted with .57 for the subarea method.

Again, one can set P and Q at preassigned small values and obtain a relationship between γ_1 and N_1. For the case $P = Q = .05$ this expression is different from that obtained in Section 7.

Thus, when $P = Q = .05$ one has:

$$(r_2 - 1)\sqrt{N_1} = 1.95996$$

Also $\dfrac{r_2 - \gamma_1}{\sqrt{\gamma_1}}\sqrt{N_1} \approx -1.64485^{+}$ for $\gamma_1 > 1$

and $\dfrac{(-r_2 - \gamma_1 + 2)\sqrt{N_1}}{\sqrt{\gamma_1}} \approx 1.64485^{+}$ for $\gamma_1 < 1$

Eliminating r_2 from these equations and solving for γ_1 we obtain

$$\gamma_1 = \frac{2\sqrt{N_1}[c_1\,\mathrm{sgn}(\gamma_1 - 1) + \sqrt{N_1}] + c_2^2 + c_2\,\mathrm{sgn}(\gamma_1 - 1)\sqrt{4\sqrt{N_1}[c_1\,\mathrm{sgn}(\gamma_1 - 1) + \sqrt{N_1}] + c_2^2}}{2N_1}$$

for $\gamma_1 > 0$

where $c_1 = 1.95996$

$c_2 = 1.64485^{+}$

Figure 15 gives the graph of N_1 vs γ_1 corresponding to the above equation. As in the previous section this value of N_1 (for a given value of γ_1) is the minimum expected number of grains for the test area which is required such that neither of the probabilities of Type I and Type II errors exceeds .05 (for the whole area method). The values of γ_1 corresponding to various values of N_1 from 5 to 1000 is given in the table below.

N_1	γ_1 $(\gamma_1 > 1)$	γ_1 $(\gamma_1 < 1)$
5	3.190	.0198*
10	2.431	.1674*
20	1.952	.3455+
30	1.756	.4424
40	1.643	.5052
50	1.569	.5503
100	1.390	.6694
200	1.270	.7600
500	1.167	.8447
1000	1.117	.8890

As an example, the graph in Figure 15 yields $N_1 = 18$ for $\gamma_1 = .31$ and 2.1. Corresponding values of N_2 for these values of N_1 are respectively 5.6 and 37.8. Thus if it is desired to distinguish between the test area and surrounding areas containing either 0.31 times as many grains per area (on the average) as the test area, or 2.1 times as many grains per area (on the average) as the test area, then the minimum required number of grains in the test area (on the average) is 18 if neither of the probabilities, P and Q, of Type I and Type II errors is to exceed .05. As another example of the use of Figure 5, suppose it is

*These figures are not accurate to four places since $N_2 < 5$ for these cases.

desired to distinguish between the test area and surrounding areas containing either 0.5 times as many grains per area or 2.0 times as many grains per area as the test area. Then $\gamma_1 = 0.5$ yields $N_1 = 40$ and $\gamma_1 = 2.0$ yields $N_1 = 18$ in the graph. The larger of the two values for N_1, namely 40, then is the minimum required number of grains in the test area if neither of the probabilities, P and Q, of Type I and Type II errors is to exceed .05.

To obtain 100(1-P)% confidence intervals for N_1, plot r_1N_1, r_2N_1 versus N_1 for a pre-assigned value of P. Then for a given value of n (on the r_1N_1, r_2N_1 axis) one obtains 100(1-P)% lower and upper confidence limits for N_1, L_{N_1} and U_{N_1} (on the N_1 axis), from this graph. Figure 16 gives the graphs of n versus (L_{N_1}, U_N) for P=.10 .05 .01. For example, suppose that 10 points are counted in the test area. Thus n = 10. For P = .05, the corresponding values for L_{N_1} and U_{N_1} are 7 and 19 respectively. This means that if we count 10 points in the test area we can expect (with a 95% chance of being right in the long run) the average number of grains in the test area to be between 7 and 19.

Conclusion

This report analyzes procedures for the evaluation of strongly enlarged photographs by electronic scanning techniques. The enlargement is so great that the grain structure of the photograph limits the accuracy of the evaluation. Two sources of uncertainty are present in this process: (1) The quantity to be determined for a certain point of the photograph (in many cases the local energy density of the electromagnetic waves) must be approximated by an integral over an area containing this point; (2) In the determination of this integral an uncertainty occurs because of the randomness in the grain distribution. To reduce the first uncertainty the test area should be taken small; to reduce the second one the test area must be taken large.

Regarding the second source of errors it is shown how the variance of the results can be expressed by a integral containing the light density and a weight function. By a suitable choice of the weight function the variance is minimized. This procedure is illustrated in several examples. Included is the problem of scanning for special patterns in the light density. The uncertainty mentioned under (1) is connected with the size of the test area and the properties of the photograph. In principle the combined variance due to the two sources of error can be minimized.

Further investigations are concerned with the discrimination between two known light densities and a special stepwise procedure in evaluating the light density of the test area.

Formulae for the confidence limits on the true grain count are given for different cases. The confidence limits depend naturally upon the average number of grains occurring in the test area. (If several photographs of the same object exist, the number of points in a given area are added). For moderate accuracy requirements (5% to 10% uncertainty) a lower limit is given by 1000 grains.

APPENDIX I

DERIVATION OF FORMULAS FOR M_Y AND V_Y

Uniform Distribution of Grains in an Area

The statistical theory given in this appendix is not self-contained. For the convenience of the reader, references have been made to one book only, "The Advanced Theory of Statistics, Vol. 1" by M. G. Kendall, [1].

Let us assume that a constant number of grains, N, are distributed with uniform probability in an area γ. Further, assume that grains arising in some subarea $\Delta\omega$ of γ are statistically independent phenomena. One may proceed by raising the following simple question: If one counts u grains in an area $\Delta\omega$ what are the limits of error on this number u?

Since the production of grains occurs with uniform probability, it follows that

$$p = \frac{\Delta\omega}{\gamma} \tag{A1}$$

is the probability that a grain is found in $\Delta\omega$, provided that we know that it is in γ.

Let X represent the number of grains counted in $\Delta\omega$. It will vary according to the particular photograph.* Since p is a constant for an area of size $\Delta\omega$ in γ it follows that the probability that X is equal to a number u (written as $P_X(u)$), is given by a

*Statisticians refer to it as a "random variable." ([1] page 173) binomial distribution

$$P_X(u) = C_u^N \, p^u (1-p)^{N-u} \qquad u = 0, 1 \ldots N \qquad \text{(A2)}$$

For a discussion of the binomial distribution see [1] page 116. In general, for our problem N will be very large and $p \ll 1$. Let $\lambda = N_p$ represent the average number of grains expected in $\Delta\omega$. If $N \to \infty$ and $\Delta\omega \to 0$ in such a manner that λ is a constant less than 5, $P_X(u)$ approaches a Poisson distribution with mean λ ; i. e.

$$P_X(u) = \frac{e^{-\lambda} \lambda^u}{u!} \qquad \text{(A3)}$$

It is well known that for the Poisson distribution the mean and the variance are both equal to λ . For a discussion of this distribution see [1] p. 120.

At this point, it is worthwhile to introduce the notion of the "characteristic function" $\phi_X(t)$ which is uniquely associated with $P_X(u)$. The introduction of $\phi_X(t)$ facilitates greatly the derivation of formulas for M_Y and V_Y (see page 4 of this report for definitions of M_Y and V_Y). $\phi_X(t)$ is defined by

$$\phi_X(t) = \sum_{u=0}^{\infty} e^{itu} P_X(u) \qquad \text{(A4)}$$

$\phi_X(t)$ is known in statistics as the "characteristic function of $P_X(u)$ (corresponding to the random variable X). See [1] page 90 for further discussion on this subject. Using (A3) and (A4) we obtain

$$\phi_X(t) = \sum_{u=0}^{\infty} e^{itu} \frac{e^{-\lambda}\lambda^u}{u!} = e^{-\lambda} \sum_{u=0}^{\infty} \frac{(\lambda e^{it})^u}{u!} = e^{-\lambda} e^{\lambda e^{it}} \tag{A5}$$

$$= e^{\lambda(e^{it}-1)}$$

Now consider γ as being subdivided into a finely divided grid of squares where each square has the same area $\Delta\omega$. Label the squares of the grid as $\Delta\omega_1 \ldots \Delta\omega_i \ldots \Delta\omega_n$ (Fig. 17).

We now consider a random variable Y

$$Y = \sum_{i=1}^{n} \theta_i X_i \tag{A6}$$

where X_i = number of grains counted in $\Delta\omega_i$

θ_i = weight assigned to the points in $\Delta\omega_i$.

It is assumed that $X_1 \ldots X_n$ are statistically independent of one another. The corresponding characteristic function of Y is obtained as follows:

First one finds

$$\phi_{\theta_i X_i}(t) = \sum_{u_i=0}^{\infty} e^{it(\theta_i u_i)} P(u_i) = e^{\lambda(e^{it\theta_i}-1)} \tag{A 7}$$

The X_i's are mutually independent. Hence for a fixed set of values $u_1, \ldots u_n$ of $X_1 \ldots X_n$ respectively we have:

$$P_Y\left(\sum_{i=1}^{u} \theta_i u_i\right) = P(u_1, u_2, \ldots, u_n) \tag{A8}$$

$$= \prod_{i=1}^{n} P(u_i)$$

See [1] page 21.

Hence

$$\phi_Y(t) = \sum_{u_1=0}^{\infty} \ldots . \sum_{u_n=0}^{\infty} e^{it(\theta_1 u_1 + \ldots + \theta_n u_n)} P_Y\left(\sum_{i=1}^{n} \theta_i u_i\right)$$

$$= \sum_{u_1=0}^{\infty} \ldots . \sum_{u_n=0}^{\infty} e^{it(\theta_1 u_1 \ldots + \theta_n u_n)} \prod_{i=1}^{n} P(u_i)$$

$$= \sum_{u_1=0}^{\infty} e^{it\theta_1 u_1} P(u_1) \sum_{u_2=0}^{\infty} e^{it\theta_2 u_2} P(u_2) \ldots \sum_{u_n=0}^{\infty} e^{itu_n} P(u_n) \tag{A9}$$

$$= \prod_{i=0}^{n} \phi_{\theta_i X_i}(t) \qquad \text{see [1] p242}$$

(A 7) and (A9) yield

$$\phi_Y(t) = \prod_{i=1}^{n} e^{\lambda(e^{it\theta_i} - 1)} = e^{\lambda \sum_{i=1}^{n} (e^{it\theta_i} - 1)} \tag{A10}$$

If we take $\log \phi_Y(t)$ and expand it in terms involving $\frac{(it)^r}{r!}$,
then the coefficient of $(it)^r/r!$ is called the r-th cumulant,
denoted by K_r. One has

$$\log \phi_Y(t) = \lambda \sum_{i=1}^{n} (e^{it\theta_i} - 1)$$

$$= \lambda \sum_{i=1}^{n} [\sum_{r=1}^{\infty} \frac{(it\theta_i)^r}{r!}] = \sum_{r=1}^{\infty} \lambda \sum_{i=1}^{n} (\theta_i)^r \frac{(it)^r}{r!}$$

Hence

$$K_r = \lambda \sum_{i=1}^{n} \theta_i^r \tag{A11}$$

It can be shown, see [1] page 60-63, that K_1 is the mean M_y of Y and that K_2 is the variance V_y of Y.

Hence

$$M_Y = \lambda \sum_{i=1}^{n} \theta_i$$

$$V_Y = \lambda \sum_{i=1}^{n} \theta_i^2 , \tag{A12}$$

where

$$\lambda = N \frac{\Delta\omega}{\gamma} \tag{A13}$$

Now introduce rectangular coordinates (x, y) and let $\Delta\omega \rightarrow 0$
Then $\theta_i = \theta(x, y)$ and $\phi_Y(t)$ becomes

$$\phi_Y(t) = \exp\left\{\frac{N}{\Upsilon}\sum_{i=1}^{n}\Delta\omega_i\left(e^{it\theta_i}-1\right)\right\} \rightarrow \exp\left\{\frac{N}{\Upsilon}\iint_{\Upsilon}\left[e^{it\theta(x,y)}-1\right]dxdy\right\} \qquad (A14)$$

assuming that the double integral exists.

From this one obtains

$$\begin{aligned} \log\phi_Y(t) &= \frac{N}{\Upsilon}\iint_{\Upsilon}\left[e^{it\theta(x,y)}-1\right]dxdy \\ &= \frac{N}{\Upsilon}\iint_{\Upsilon}\sum_{r=1}^{\infty}\frac{[it\theta(x,y)]^r}{r!}\,dxdy \\ &= \sum_{r=1}^{\infty}\frac{(it)^r}{r!}\left[\frac{N}{\Upsilon}\iint_{\Upsilon}\theta^r(x,y)\,dxdy\right] \end{aligned} \qquad (A15)$$

This yields

$$K_r = \frac{N}{\Upsilon}\iint_{\Upsilon}\theta^r(x,y)\,dxdy \qquad (A16)$$

Thus

$$M_y = \frac{N}{\Upsilon}\iint_{\Upsilon}\theta(x,y)\,dxdy \qquad (A17)$$

$$V_Y = \frac{N}{\Upsilon}\iint_{\Upsilon}\theta^2(x,y)\,dxdy \qquad (A18)$$

General Case

It has been assumed in the text that the grains arise by a random process in such a manner that the probability that a grain is produced in a given area is proportional to the light flux across that area. On the basis of this assumption, it is natural to make the following definition for p, where p is the probability of a grain in $\Delta\omega_i$ given that the grain is in γ. Define p as:

$$p = \frac{\iint_{\Delta\omega_i} \rho(x, y)\, dxdy}{\iint_{\gamma} \rho(x, y)\, dxdy} \tag{A19}$$

where $\rho(x, y)$ is the density of the light flux.

Then, since $N = k \iint_{\gamma} \rho(x, y)\, dxdy$ by Equation 1

$$p = \frac{k}{N} \iint_{\Delta\omega_i} \rho(x, y)\, dxdy \tag{A20}$$

Note that when $\rho(x, y)$ is a constant and $\Delta\omega_i$ has area $\Delta\omega$, $p = \frac{\Delta\omega}{\gamma}$, and this reduces to the case discussed in Section 1.

We shall assume that the region γ has been subdivided by some suitable method into subdivisions $\Delta\omega_i$ such that p, as defined in (A19), is the same for any $\Delta\omega_i$ thus produced. Then the discussion in Section 1 can be applied again. Here, $\lambda = Np = k \iint_{\Delta\omega_i} \rho(x, y) dxdy$ is the average number of grains expected in $\Delta\omega_i$. By definition this is a constant for any i.

One obtains in a manner similar to that given in Section 1,

$$\log \phi_Y(t) = k \sum_{i=1}^{n} \left[\iint_{\Delta\omega_i} \rho(x, y)\, dxdy \right] \left[e^{it\theta_i} - 1 \right]$$

or

$$K_r = k \sum_{i=1}^{n} \left[\iint_{\Delta\omega_i} \rho(x, y)\, dxdy \right] \theta_i^r \qquad \text{(A21)}$$

If one considers θ_i as a discontinuous function of x and y which assumes in each region $\Delta\omega_i$ the value θ_i defined above, one can write the last expression

$$K_r = k \iint_Y \rho(x, y)\, \theta_i^r(x, y)\, dxdy$$

If $\Delta\omega_i \to 0$, $\theta_i(x, y)$ tends to a function $\theta(x, y)$, i. e.

$$[\theta_i(x, y) - \theta(x, y)] = 0$$

Then

$$K_r = k \iint_Y \rho(x, y)\, \theta^r(x, y)\, dxdy$$

Hence

$$M_Y = k \iint_Y \rho(x, y)\, \theta(x, y)\, dxdy \qquad \text{(A22)}$$

$$V_Y = k \iint_Y \rho(x, y)\, \theta^2(x, y)\, dxdy \qquad \text{(A23)}$$

REFERENCE:

1. Kendall, M. G., The Advanced Theory of Statistics, Vol. I, 1948, London, Charles Griffin and Company, Limited.

APPENDIX II

COMPARISON OF POISSON DISTRIBUTION AND NORMAL DISTRIBUTION

The assumption that the grains of the photograph arise in a random fashion and that the creation of a new grain can be considered as an event independent of the creation of all previous grains leads to a Poisson distribution. To be specific if the average number of grains to be found in a certain test area is N then the probability of finding in such a test area n grains is given by

$$p(n) = e^{-N} \frac{N^n}{n!} \tag{A24}$$

For large values of n this expression can be approximated by a normal distribution. It is obtained if one expresses n! by means of Stirling's formula and then carries out certain simplifications, using the assumption that the number "n" differs from the average value N by only a small amount. One thus obtains $\bar{p}_n$, the probability density of finding n grains in a given area of size dn(if the expected number is N) as the normal law:

$$\bar{p}(n) = \frac{1}{\sqrt{2\pi N}} e^{-\frac{(n-N+\frac{1}{2})^2}{2N}} \tag{A25}$$

In Figs. 18-23 the expressions (A24) and (A25) are plotted for N = 10 (10) 60. One recognizes that the normal distribution gives a good approximation for N as low as 10 (In fact it is fairly good for

N as low as 5). Thus it appears to be justified if one takes the normal distribution as approximation for the Poisson distribution whenever such an approximation is needed.

Table I

Values of P and P_s for the case when $r' = r$ $(\gamma_1 > 1)$

$(r-1)\sqrt{N'_1}$	p_s	P_s	P	P/P_s
0	.5000	.5000	.5000	1.000
0.1	.4602	.4028	.3821	.949
0.2	.4207	.3113	.2743	.881
0.3	.3821	.2301	.1841	.800
0.4	.3446	.1628	.1151	.707
0.5	.3085	.1097	.0668	.609
0.6	.2743	.0704	.0359	.510
0.7	.2420	.0429	.0179	.416
0.8	.2119	.0250	.0082	.328
0.9	.1841	.0138	$.003\bar{5}$	.251
1.0	.1587	.0072	.0014	.187
1.1	.1357	.0036	$.000\bar{5}$	.134
1.2	.1151	.0017	.0002	.094

Table II(a)

Values of r for $\gamma_1 > 1$

N_1	$P = P_s = .10$	$P = P_s = .05$	$P = P_s = .01$
5	1.573	1.736	2.040
10	1.405	1.520	1.736
20	1.287	1.368	1.520
30	1.234	1.300	$1.42\bar{5}$
40	1.203	1.260	1.368
50	1.181	1.233	1.329
100	1.128	1.164	1.233
200	1.091	1.116	1.164
500	1.057	1.074	1.104
1000	1.041	1.052	1.074

Table II(b)

Values of r' for $\gamma_1 > 1$

N_1	$P = P_s = .10$	$P = P_s = .05$	$P = P_s = .01$
50	1.221	1.284	1.403
100	1.156	1.201	$1.28\overset{+}{5}$
200	1.111	1.142	1.202
500	1.070	1.090	1.128
1000	1.049	1.064	1.090

Table III

Values of Q and Q_s for $\gamma_1 > 1$

$P = P_s = .10$

	$N_1 = 50$		$N_1 = 100$		$N_1 = 200$		$N_1 = 500$		$N_1 = 1000$	
γ_1	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q
1.1	.7488	.7081	.6706	.6058	.5469	.4497	.2998	.1814	.1061	.0365
1.2	.5449	.4518	.3723	.2560	.1722	.0790	.0151	.0018	.0002	.0000
1.3	.3445	.2307	.1512	.0659	.0275	.0047	.0001	.0000	.0000	.0000
1.5	.0938	.0329	.0110	.0012	.0001	.0000	.0000	.0000	.0000	.0000
2.0	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

$P = P_s = .05$

	$N_1 = 50$		$N_1 = 100$		$N_1 = 200$		$N_1 = 500$		$N_1 = 1000$	
γ_1	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q
1.1	.8458	.8144	.7852	.7307	.6795	.5870	.4300	.2865	.1840	.0740
1.2	.6724	.5834	.5032	.3729	.2702	.1400	.0331	.0049	.0007	.0000
1.3	.4682	.3380	.2384	.1173	.0546	.0114	.0004	.0000	.0000	.0000
1.5	.1538	.0613	.0231	.0031	.0004	.0000	.0000	.0000	.0000	.0000
2.0	.0018	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

$P = P_s = .01$

	$N_1 = 50$		$N_1 = 100$		$N_1 = 200$		$N_1 = 500$		$N_1 = 1000$	
γ_1	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q
1.1	.9528	.9387	.9257	.8970	.8692	.8078	.6847	.5343	.4039	.2127
1.2	.8588	.7975	.7380	.6171	.5068	.3234	.1127	.0251	.0049	.0001
1.3	.7002	.5714	.4573	.2773	.1586	.0464	.0029	.0001	.0000	.0000
1.5	.3234	.1617	.0755	.0145	.0025	.0001	.0000	.0000		.0000
2.0	.0075	.0004	.0000	.0000	.0000	.0000	.0000	.0000		.0000

Table IV(a)

Values of r_1 and r_2 for $\gamma_1 \neq 1$

	$P=P_s=.10$		$P=P_s=.05$		$P=P_s=.01$	
N_1	r_1	r_2	r_1	r_2	r_1	r_2
5	.264	1.736	.123	1.877	-.152	2.152
10	.480	1.520	.380	1.620	.185	1.815
20	.632	1.368	.562	1.438	.424	1.576
30	.700	1.300	.642	1.358	.530	1.470
40	.740	1.260	.690	1.310	.593	1.407
50	.767	1.233	.723	1.277	.636	1.364
100	.836	1.164	.804	1.196	.742	1.258
200	.884	1.116	.861	1.139	.818	1.182
500	.926	1.074	.912	1.088	.885	1.115
1000	.948	1.052	.938	1.062	.919	1.081

Table IV(b)

Values of r_2' and r_2' for $\gamma_1 \neq 1$

	$P=P_s=.10$		$P=P_s=.05$		$P=P_s=.01$	
N_1	r_1'	r_2'	r_1'	r_2'	r_1'	r_2'
50	.561	1.439	.513	1.487	.419	1.581
100	.690	1.310	.656	1.344	.589	1.411
200	.781	1.219	.757	1.243	.710	1.290
500	.861	1.139	.846	1.154	.816	1.184
1000	.902	1.098	.891	1.109	.870	1.130

Table V

Values of Q and Q_s for $\gamma_1 \neq 1$

$P = P_s = .10$

	$N_1 = 50$		$N_1 = 100$		$N_1 = 200$		$N_1 = 500$		$N_1 = 1000$	
γ_1	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q
0.5	.3060	.0037	.0142	.0000	.0000	.0000				
0.6	.7474	$.191\overset{\pm}{5}$	.3982	.0193	.0526	.0001	.0000	.0000		
0.7	.7992	$.284\overset{+}{5}$	.5218	.0526	.1297	.0010	.0002	.0000	.0000	.0000
0.8	.8876	.6014	$.794\overset{-}{5}$	.3456	.5667	.0929	.1020	.0008	.0017	.0000
0.9	.9127	.8320	.8954	.7490	.8561	.5954	.7108	.2666	.4456	.0549
1.1	.8508	.8019	.8316	.7248	.7904	.5853	.6534	.2864	.4247	.0740
1.2	.7574	.5808	.6706	$.372\overset{-}{5}$	$.494\overset{-}{5}$	.1399	.1366	.0049	.0078	.0000
1.3	$.620\overset{+}{5}$	$.337\overset{+}{5}$	.4383	.1173	.1779	.0114	.0047	.0000	.0000	.0000
1.5	.3047	.0613	.0860	.0031	.0038	.0000	.0000	.0000		
2.0	.0089	.0001	.0000	.0000	.0000	.0000				

$P = P_s = .05$

	$N_1 = 50$		$N_1 = 100$		$N_1 = 200$		$N_1 = 500$		$N_1 = 1000$	
γ_1	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q
0.5	.4552	.0129	.0357	.0000	.0000	.0000				
0.6	.8479	.3134	.5361	.0463	.1003	.0004	.0000			
0.7	.8847	$.423\overset{+}{5}$	$.655\overset{+}{5}$	.1069	.2140	.0032	.0008	.0000	.0000	
0.8	.9427	.7290	.8791	.4821	.6916	.1658	.1697	$.002\overset{-}{5}$	.0043	.0000
0.9	.9579	.9042	.9472	.8433	$.921\overset{+}{5}$	.7173	.8127	$.385\overset{+}{5}$	.5707	$.102\overset{\pm}{5}$
1.1	.9173	.8784	.9037	.8176	.8732	.6979	.7611	.3961	.5427	.1258
1.2	.8471	.6898	.7749	.4853	.6113	.2139	.2051	.0109	.0153	.0000
1.3	.7297	.4436	.5527	.1808	.2563	.0226	.0093	.0000	.0000	
1.5	.4047	.0991	.1322	$.006\overset{\pm}{5}$	.0074	.0000	.0000			
2.0	.0156	.0002	.0001	.0000	.0000					

Table V (continued)

	$N_1 = 50$		$N_1 = 100$		$N_1 = 200$		$N_1 = 500$		$N_1 = 1000$	
γ_1	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q	Q_s	Q
0.5	.7460	.0874	.1504	.0003	.0003	.0000	.0000			
0.6	.9573	.6056	.7792	.1768	.2708	.0044	.0004	.0000		
0.7	.9703	.7065	.8574	.3061	.4462	.0232	.0056	.0000	.0000	
0.8	.9882	.9030	.9665	.7401	.8724	.3888	.3674	.0170	.0214	.0000
0.9	.9922	.9753	.9893	.9516	.9812	.8896	.9354	.6399	.7829	.2682
1.1	.9792	.9617	.9740	.9332	.9614	.8659	.9037	.6270	.7491	.2880
1.2	.9489	.8554	.9104	.7004	.8002	.4088	.3816	.0417	.0483	.0003
1.3	.8817	.6549	.7501	.3549	.4433	.0719	.0307	.0001	.0001	.0000
1.5	.6050	.2166	.2612	.0239	.0234	.0001	.0000	.0000	.0000	
2.0	.0403	.0007	.0004	.0000	.0000	.0000				

APPENDIX III

I. Computational procedure for finding r, given α an assigned value of P(x).

$$P(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}} \, du$$

where $x = (r-1)\sqrt{N_1}$

Now set P(x) equal to α, i.e.

$$P(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{u^2}{2}} \, du = \alpha \qquad \text{(a)}$$

Given α, the problem is: compute P (x). To solve this problem on Burroughs E101 Electrodata Electronic Computer, we proceed in the following manner. We replace P(x) in (a) by the approximate formula

$$P(x) \approx \left[\frac{1}{c + a_1 \bar{x} + a_2 \bar{x}^2 + a_3 \bar{x}^3 + a_4 \bar{x}^4 + a_5 \bar{x}^5 + a_6 \bar{x}^6}\right]^{16} \qquad \text{(b)}$$

where $\bar{x} = \dfrac{x}{7.0710678119}$

$c = 1.0442737826 = \dfrac{1}{\sqrt[16]{2}}$
$a_1 = .03682270091$
$a_2 = .11038499229$
$a_3 = .12101231517$
$a_4 = .00992153425$
$a_5 = .09025371127$
$a_6 = .07026624580$

for (x) in the range: $-7.0710678119 \leq x \leq 7.0710678119$

The formula (b) will yield 6-place accuracy. Differentiating (a) we obtain:

$$P'(x) = -\frac{1}{\sqrt{2\pi}} \epsilon^{-\frac{u^2}{2}} \tag{c}$$

Now, $-\frac{1}{\sqrt{2\pi}} = .398994\ 22811$

Let $x^2/2 = \bar{x}$

then

$$e^{-\bar{x}} = a_0 - a_1\bar{x} + a_2\bar{x}^2 - a_3\bar{x}^3 + a_4\bar{x}^4 - a_5\bar{x}^5 + a_6\bar{x}^6 \ldots + a_{10}\bar{x}^{10} \tag{d}$$

where

$a_0 = 1.0$
$a_1 = -1.0$
$a_2 = .5$
$a_3 = .166666\ 66667$
$a_4 = .416666\ 66667$
$a_5 = .008333\ 33333$
$a_6 = .001388\ 888889$
$a_7 = .000198\ 41270$
$a_8 = .000024\ 80159$
$a_9 = .0000027\ 5573$
$a_{10} = .0000002\ 7557$

(d) will yield 10-place accuracy if the following identity is used:

$$e^{-x} = (e^{-\frac{x}{n}})^n \quad , \text{ where } \frac{x}{n} \leq .5 \text{ for any } x > 0$$

(For example, if it is desired to compute e^{-2}, first compute $e^{-\frac{2}{4}}$, then raise to the fourth power).

Now we can compute $P(x)$ and $P'(x)$

$$\text{Let } f(x) = P(x) - \alpha$$

$$f'(x) = P'(x)$$

Iteration formula for x:

$$x_{n+1} = x_n - \frac{f(x)-\alpha}{f'(x)} = x_n - \frac{P(x)}{P'(x)} \qquad \text{(e)}$$

Take an arbitrary value for x. We calculate $P(\bar{x})$ and $P'(x)$ using (b), (c). Substituting these into (e) give us a second value for x. Repeat iteration by repeated application of (e). The process converges to the desired value of x. If this value is y, we set:

$$y = (r-1)\sqrt{n_1}$$

$$\text{or,} \quad r = (y/\sqrt{n_1}) + 1$$

Then, for a given n_1, r is computed

II. Computational procedure for finding r', given α, an assigned value of P_s where P_s is given by the formula on page 44. As before, set $P_s = \alpha$. The incomplete beta-function $I_x(p, q)$ can be represented by the following series:

$$I_x(5.5) = x^5[a_4x^4 - a_3x^3 + a_2x^2 - a_1x + a_0] \qquad \text{(f)}$$

where $a_4 = 70$

$a_3 = 315$

$a_2 = 540$

$a_1 = 420$

$a_0 = 126$

This function also represents P_s from formula given on Page 45.

Now $f(x) = \alpha - I_x(5,5)$

$$f'(x) = -630\ x^4(1-x)^4$$

To find x, we use the following iteration formula:

$$x_n + 1 = x_n - \frac{f(x)}{f'(x)}$$

$$= x_n + \frac{\alpha - I_x(5,5)}{630\ x^4(1-x^4)}$$

Using the solution of this iteration, call it y, we set:

$$y = \int_t^{\infty} e^{-\frac{u^2}{2}} \qquad du$$

where $t = (r'-1)\sqrt{N_1'}$

and the problem now reduces to solving for t. But this is exactly the same type of problem as in (I). Using the method given in (I) we compute t. Then for a given N_1 we compute r' from the relationship $u = (r'-1)\sqrt{N_1'}$.

III. Computational procedure for finding Q given γ_1, N_1 and r.

r was obtained using Computational Procedure I from page 46. Q is given by:

$$Q \approx \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_1} e^{-\frac{u^2}{2}} du$$

$$\text{where } u_1 = \frac{(r - \gamma_1)\sqrt{N_1}}{\sqrt{\gamma_1}}$$

To compute Q, assign a value to γ_1 (say, $\gamma_1 = 1.1$) and to N_1. Using the value of r obtained, we calculate u_1. With this value of u_1 we use formula (b) and calculate $P(\bar{x})$, where

$$\bar{x} = \frac{u_1}{7.0710678119}$$

This value of $P(\bar{x})$ is the desired Q.

IV. Computational Procedure for finding Q_s, given γ_1, N_1, r'. The value of r' was obtained using Computation Procedure II. Now from page 47,

$$q_s = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{u_2} e^{-\frac{u^2}{2}} du$$

$$\text{where } u_2 = \frac{(r' - \gamma_1)\sqrt{N_1}}{3\sqrt{\gamma_1}}$$

Assign a value to γ_1 (say $\gamma_1 = 1.1$) and to N_1. Using the value of r' obtained, we calculate u_2. With this value of u_2 we use formula (b) and calculate $P(\bar{x})$ where $\bar{x} = \frac{u_2}{7.0710678119}$

This value of $P(\bar{x})$ is q_s. In formula (f) set x equal to the value of q_s which we have obtained, and calculate $I_x(5, 5)$. This value is the desired value of Q_s.

V. Given the function:

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-\frac{u^2}{2}} du$$

$\Phi(x)$ can be approximated to six decimal places by the following series in the range $-7.0710678119 \leq x \leq 7.0710678119$

$$\Phi(\bar{x}) = 1 - \left[\frac{1}{a_0 + a_1\bar{x} + a_2\bar{x}^2 + a_3\bar{x}^3 + a_4\bar{x}^4 + a_5\bar{x}^5 + a_6\bar{x}^6}\right]^{16}$$

where $\bar{x} = \dfrac{x}{7.0710678119}$

$a_0 = 1$
$a_1 = .03526\ 15392$
$a_2 = .105705\ 03075$
$a_3 = .115881\ 7900$
$a_4 = .009500\ 89375$
$a_5 = .086427\ 2500$
$a_6 = .067287\ 1875$

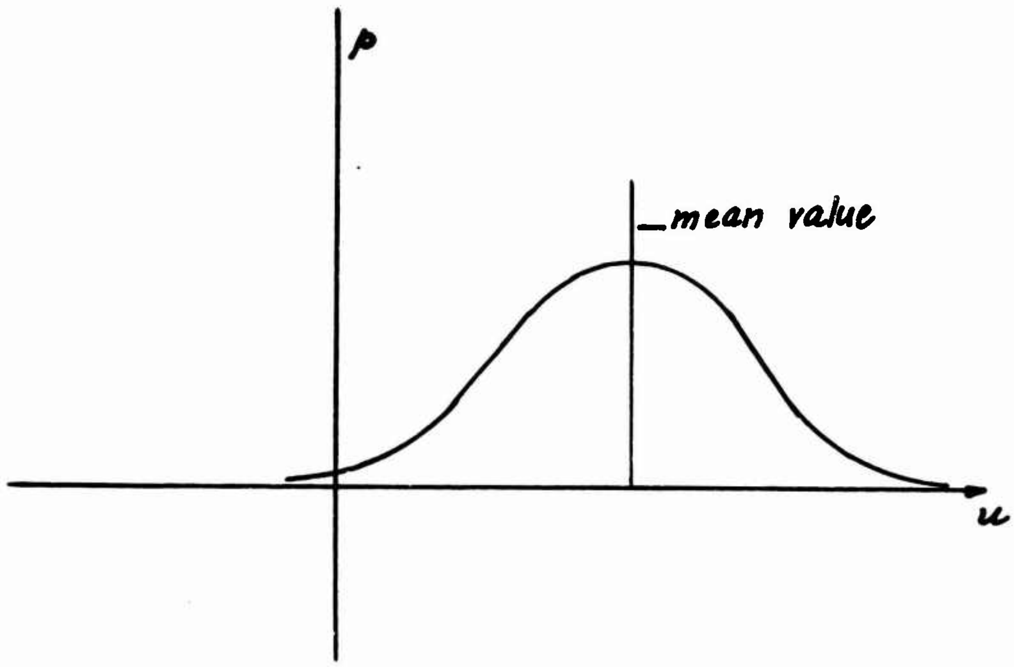

Fig. 1 PROBABILITY DENSITY CURVE

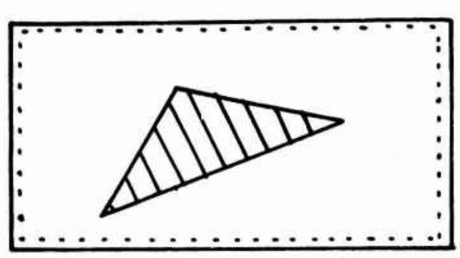

Fig. 2 SCANNING FOR A GIVEN PATTERN
THE RECTANGLE REPRESENTS THE TEST AREA A, THE SHADED PART IS THE PATTERN THAT IS TO BE FOUND. ITS AREA IS γ_0 A.

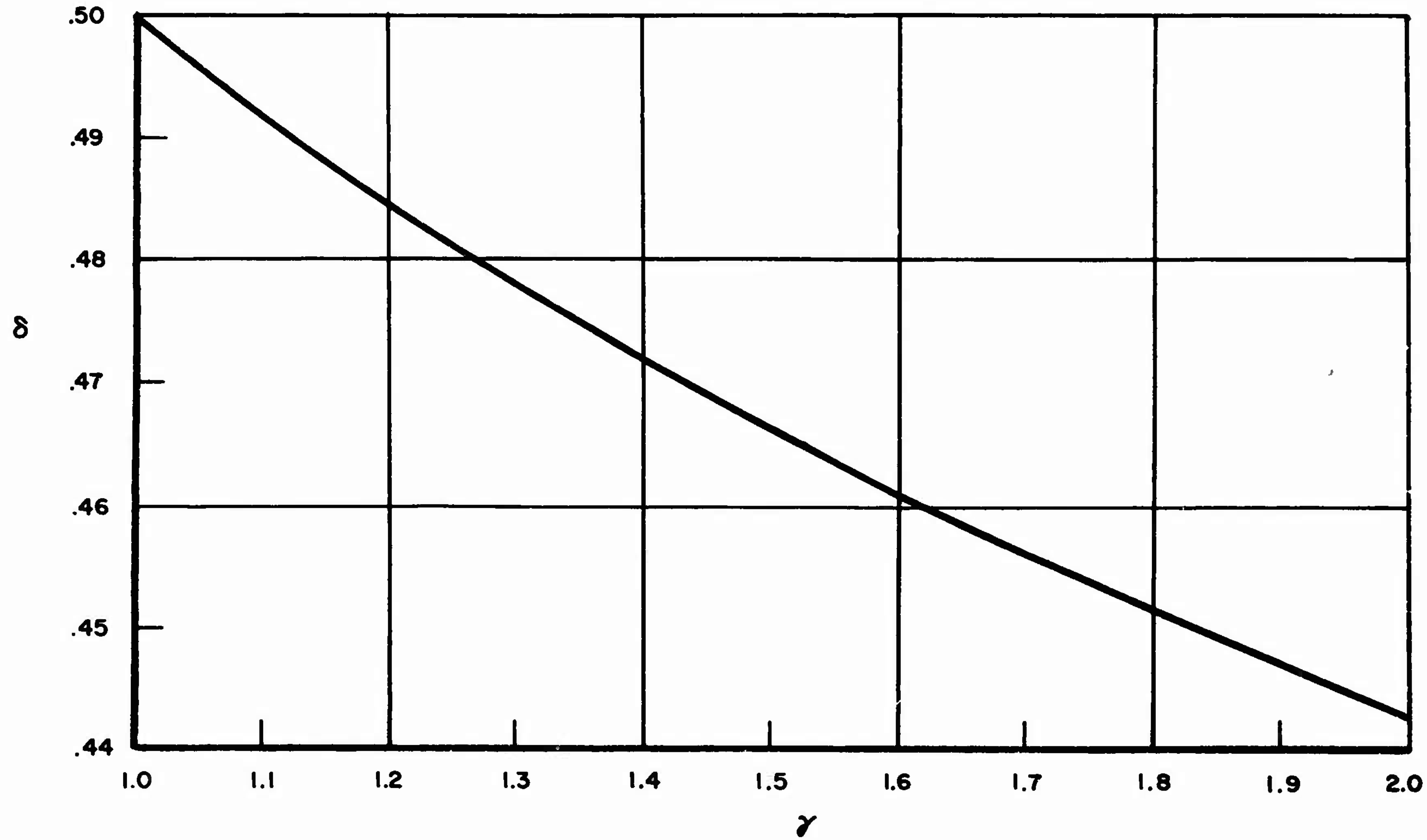

Fig. 3 GRAPH OF δ AS A FUNCTION OF γ_0

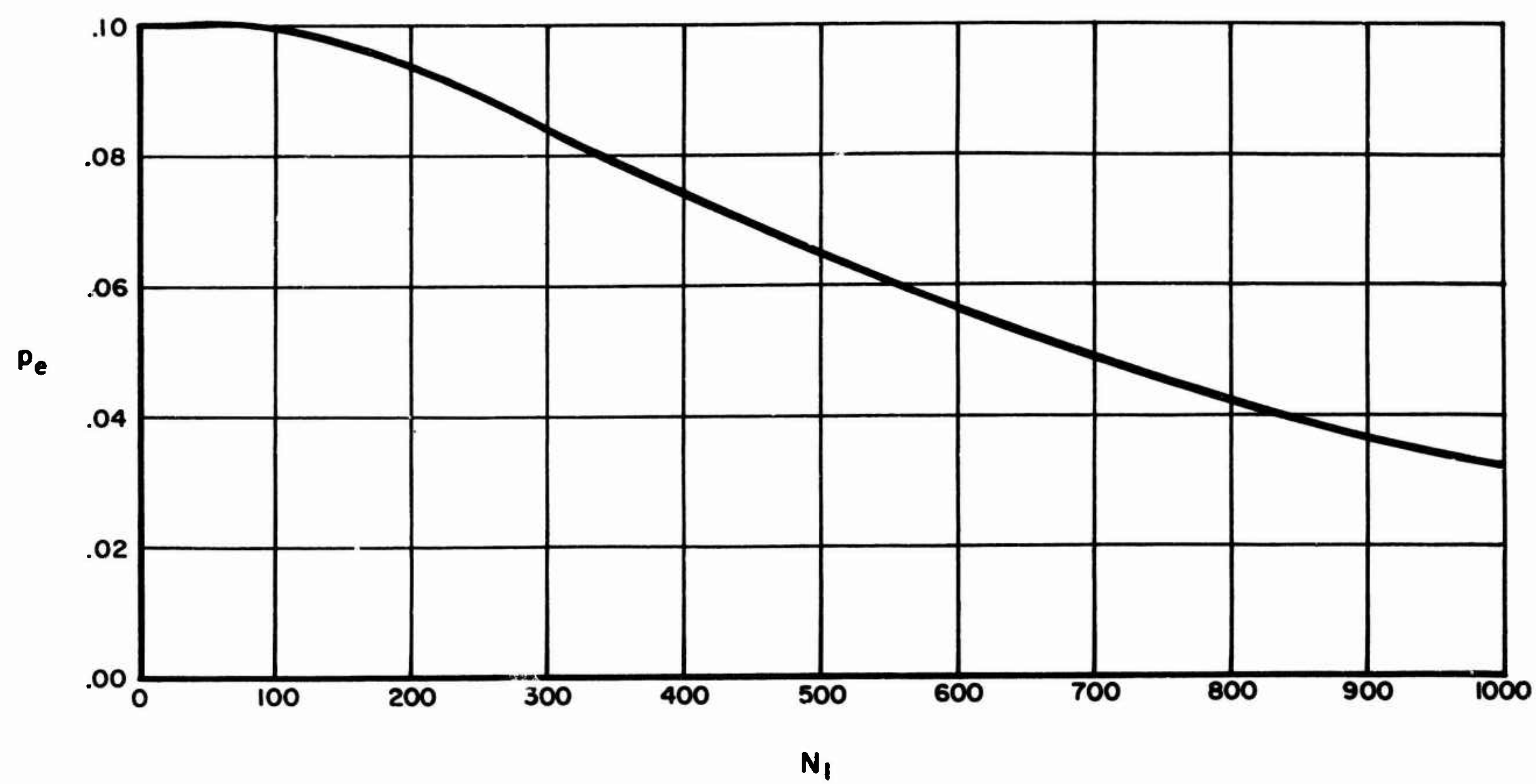

Fig. 4 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.1$, $\gamma_1 = 1.1$

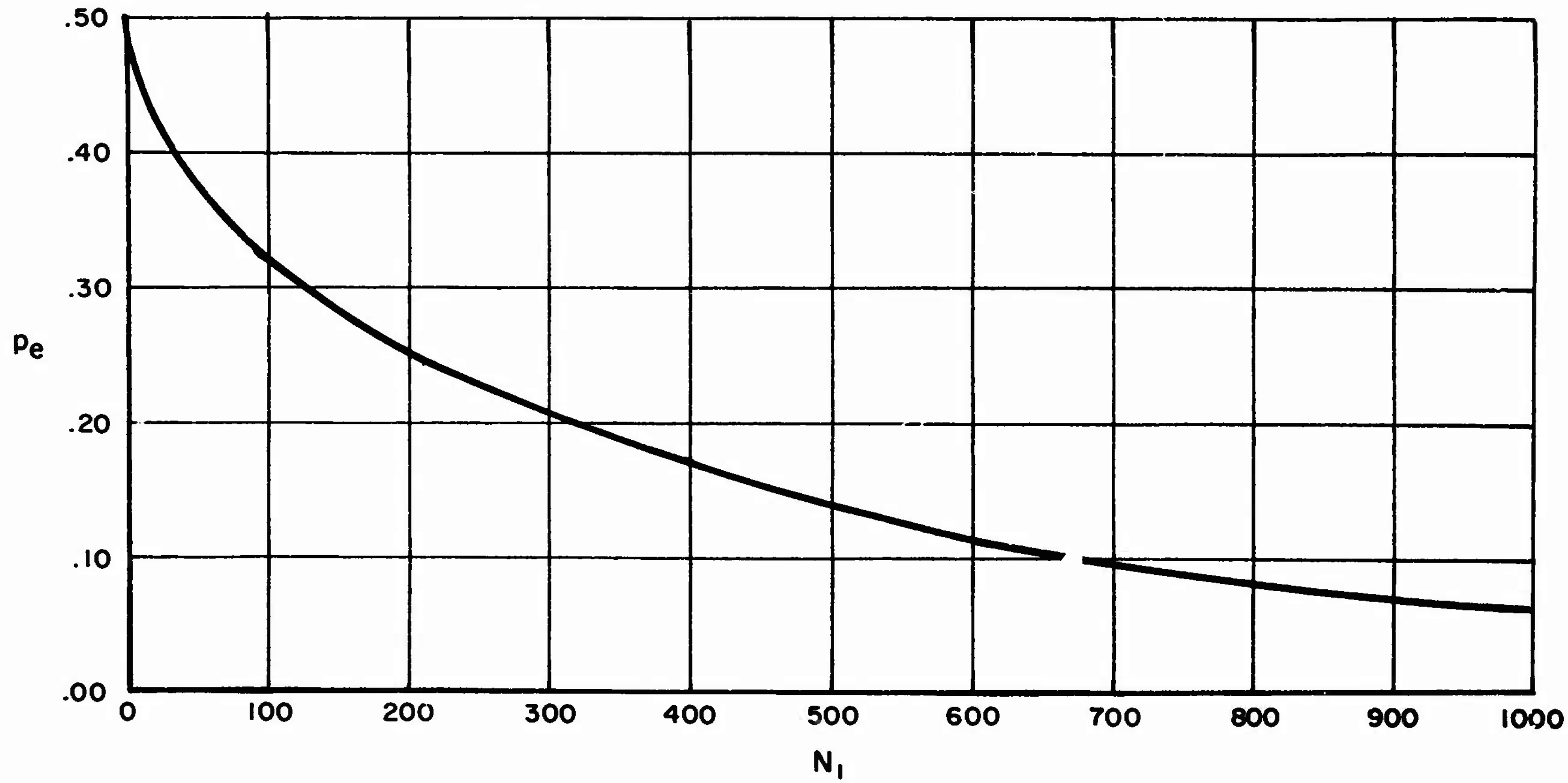

Fig. 5 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.5$, $\gamma_I = 1.1$

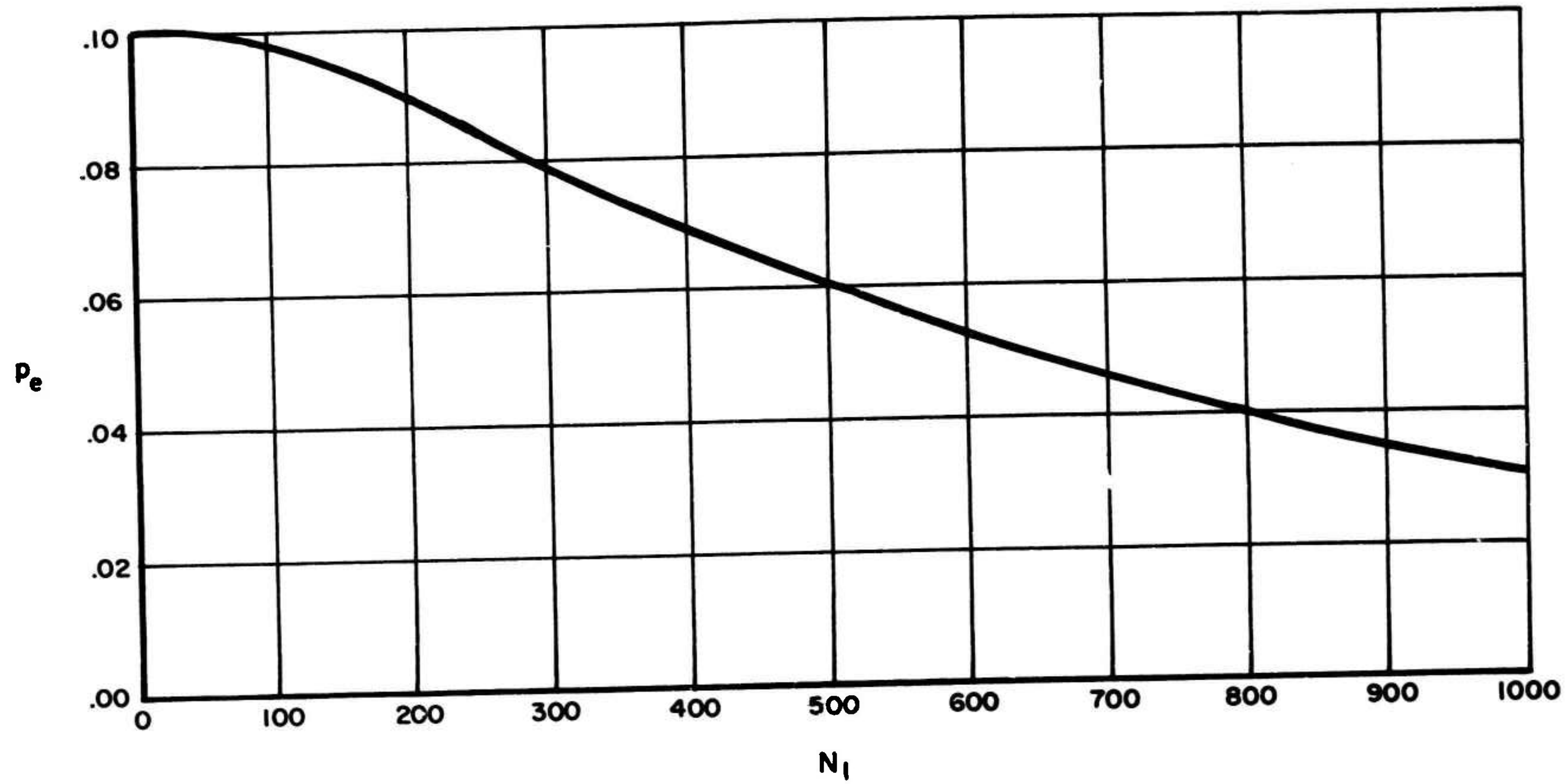

Fig. 6 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.9$, $\gamma_1 \doteq 1.1$

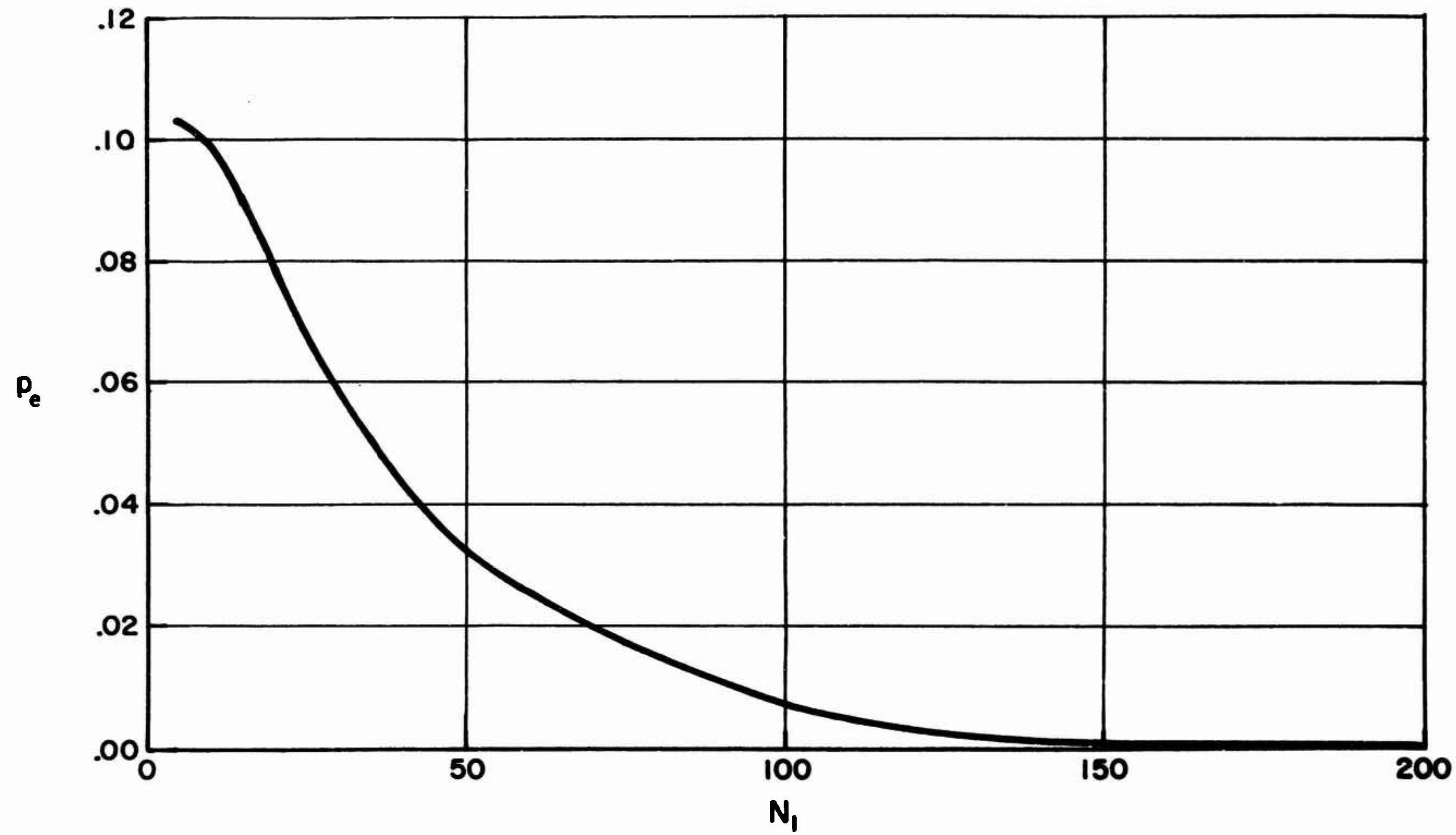

Fig. 7 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.1$, $\gamma_1 = 1.5$

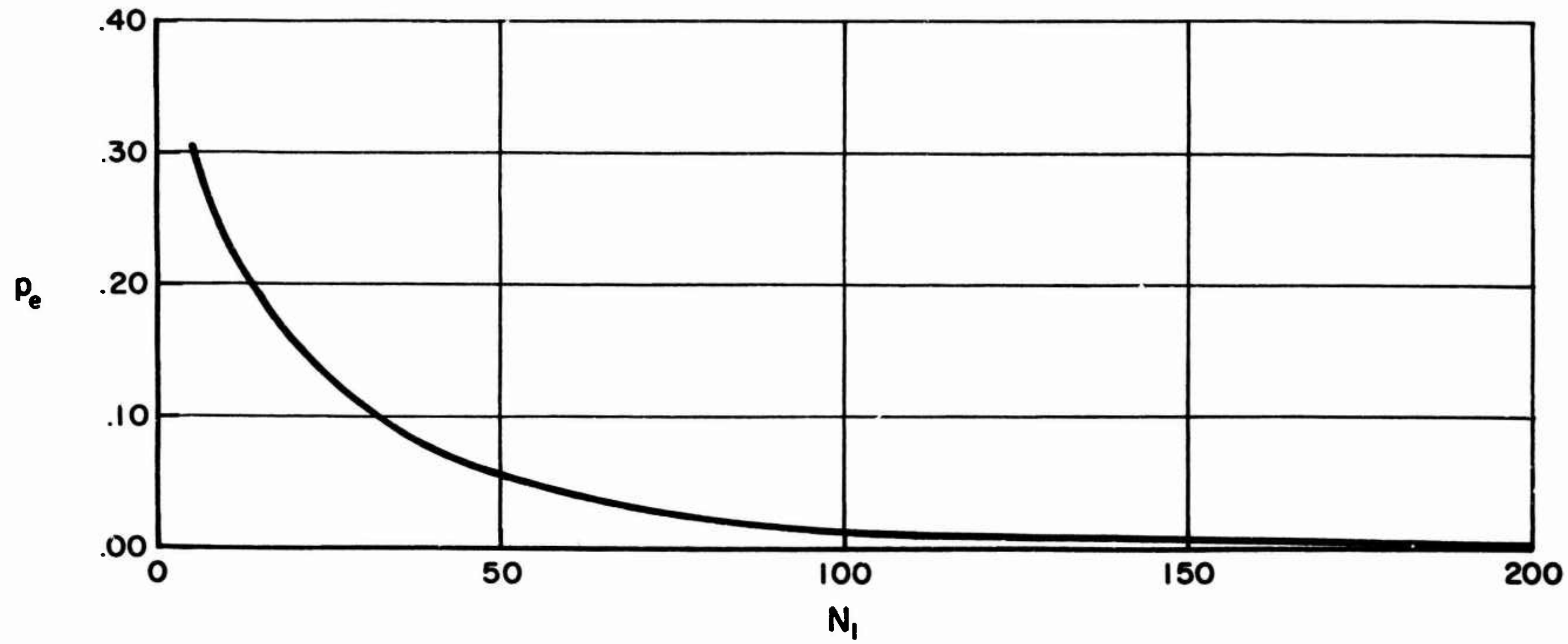

Fig. 8 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.5$, $\gamma_1 = 1.5$

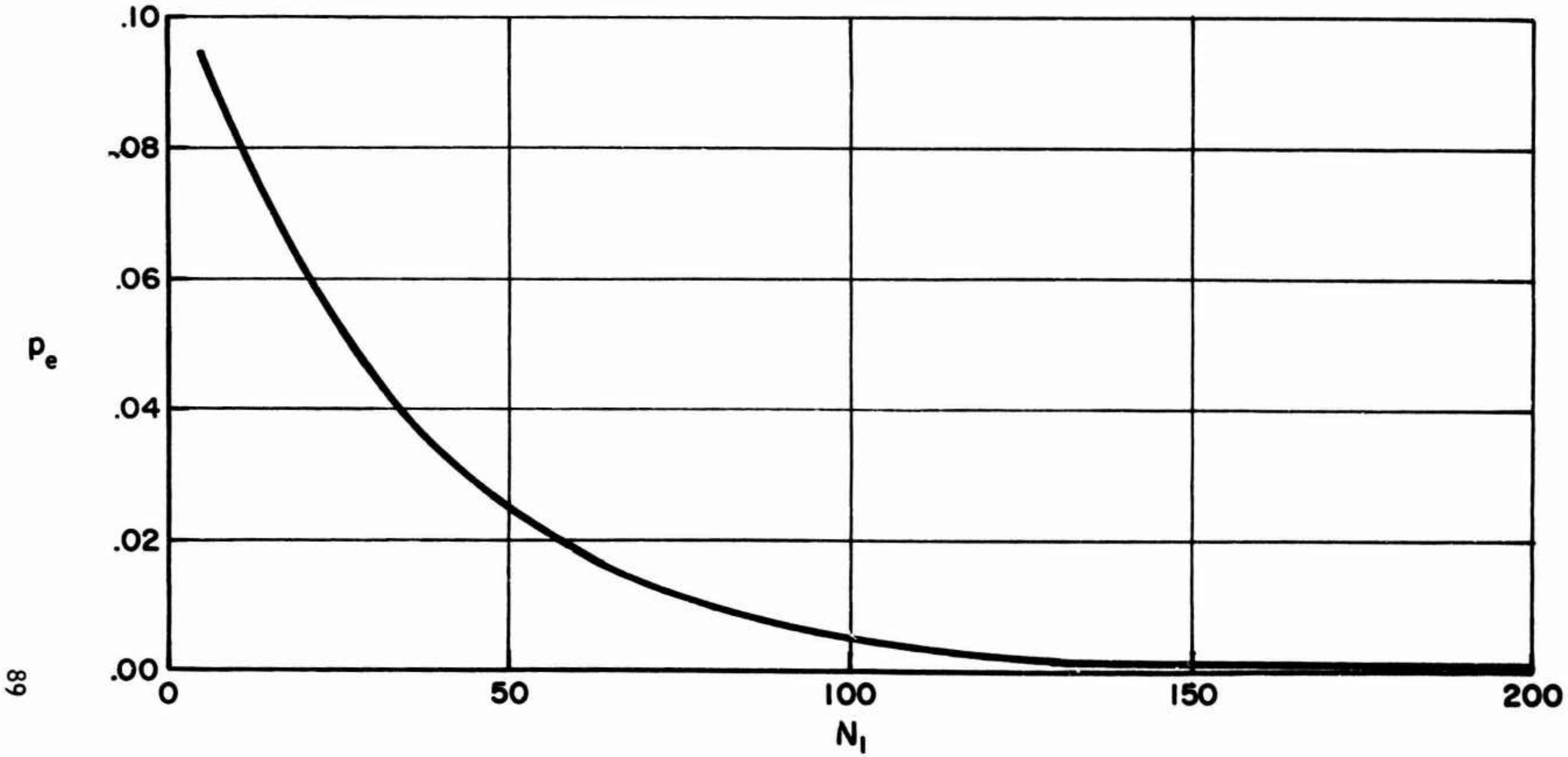

Fig. 9 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.9$, $\gamma_1 = 1.5$

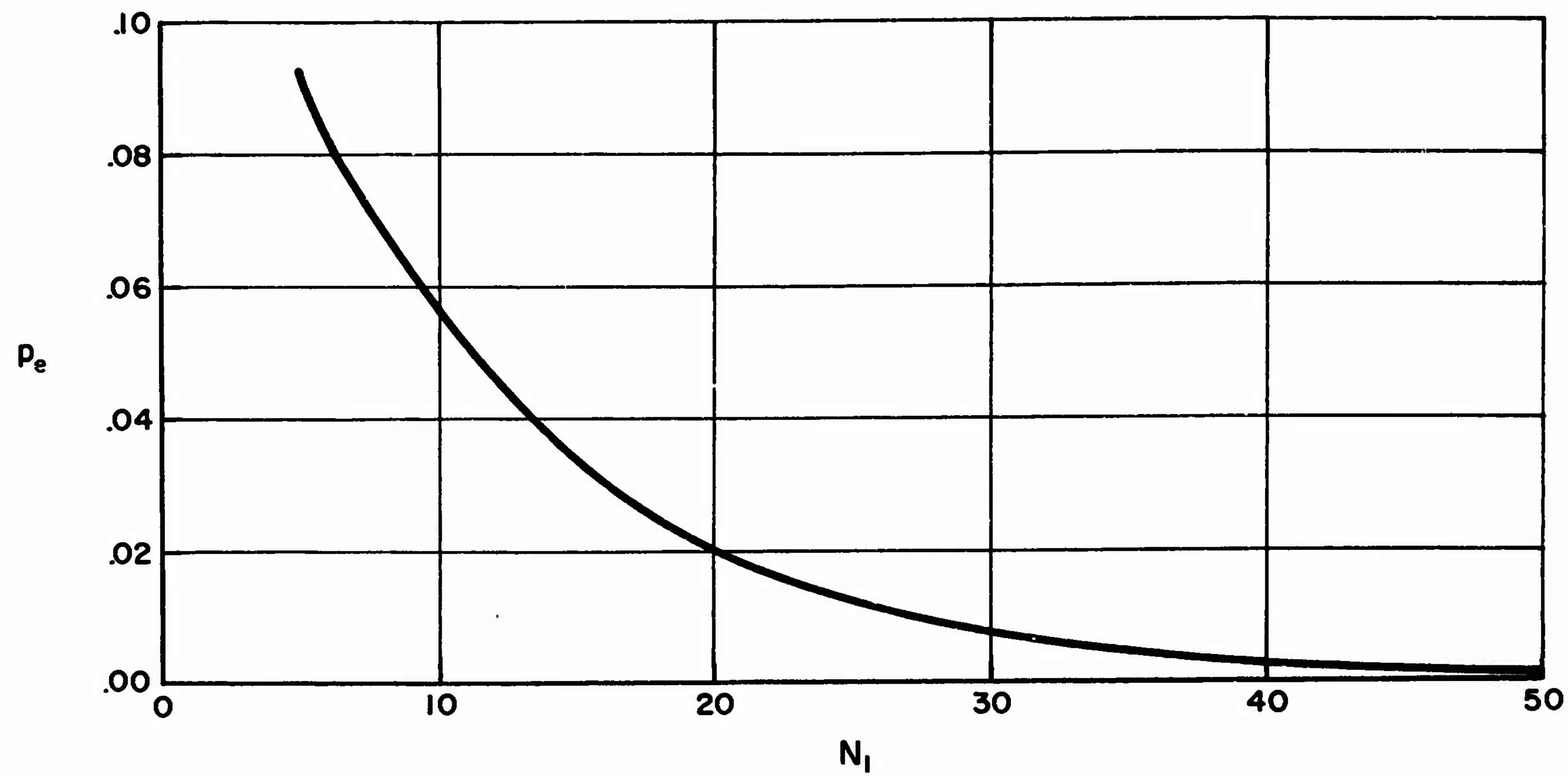

Fig. 10 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.1$, $\gamma_1 = 2.0$

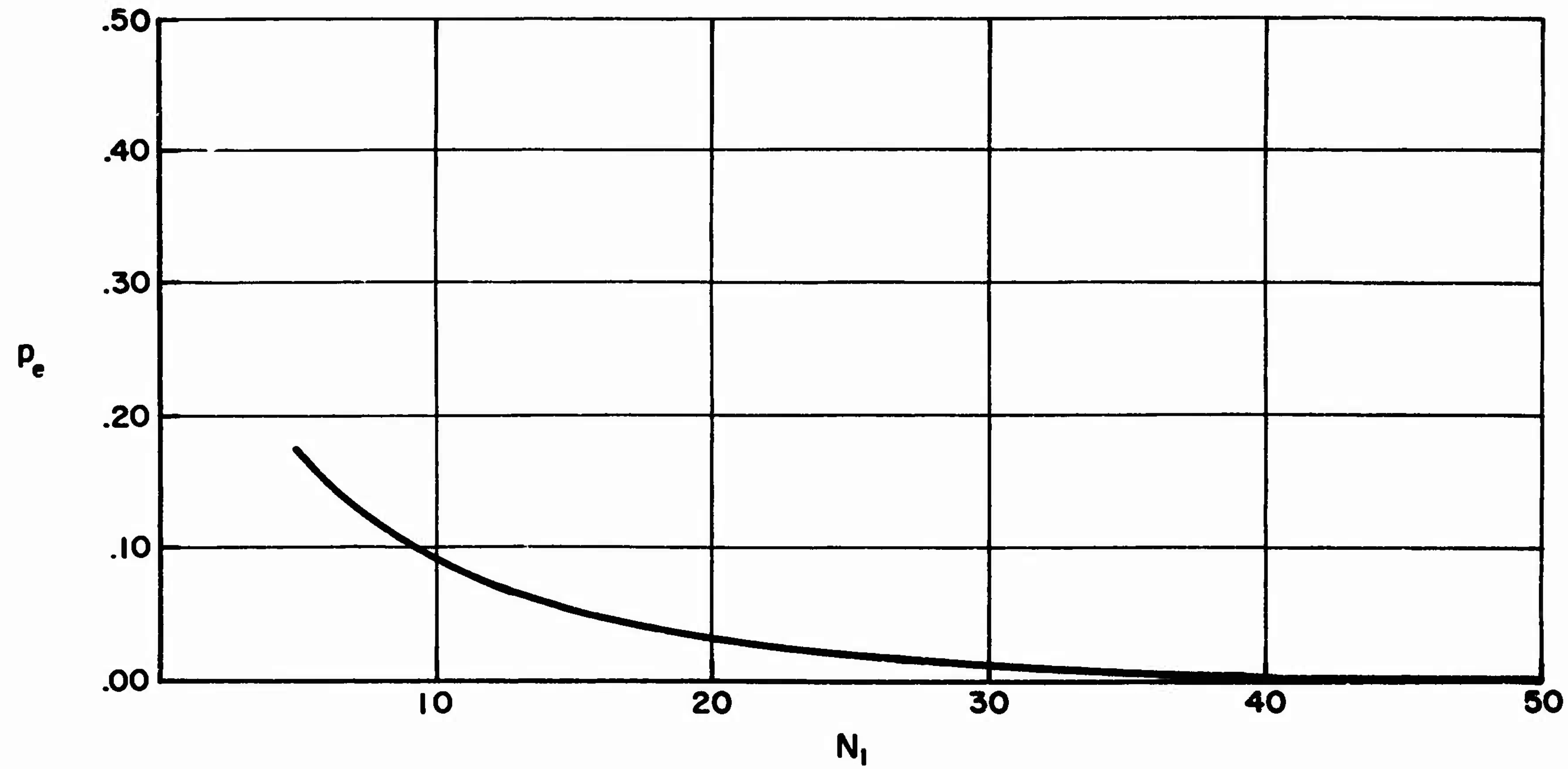

Fig. 11 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION OF N_1 FOR $\beta = 0.5$, $\gamma_I = 2.0$

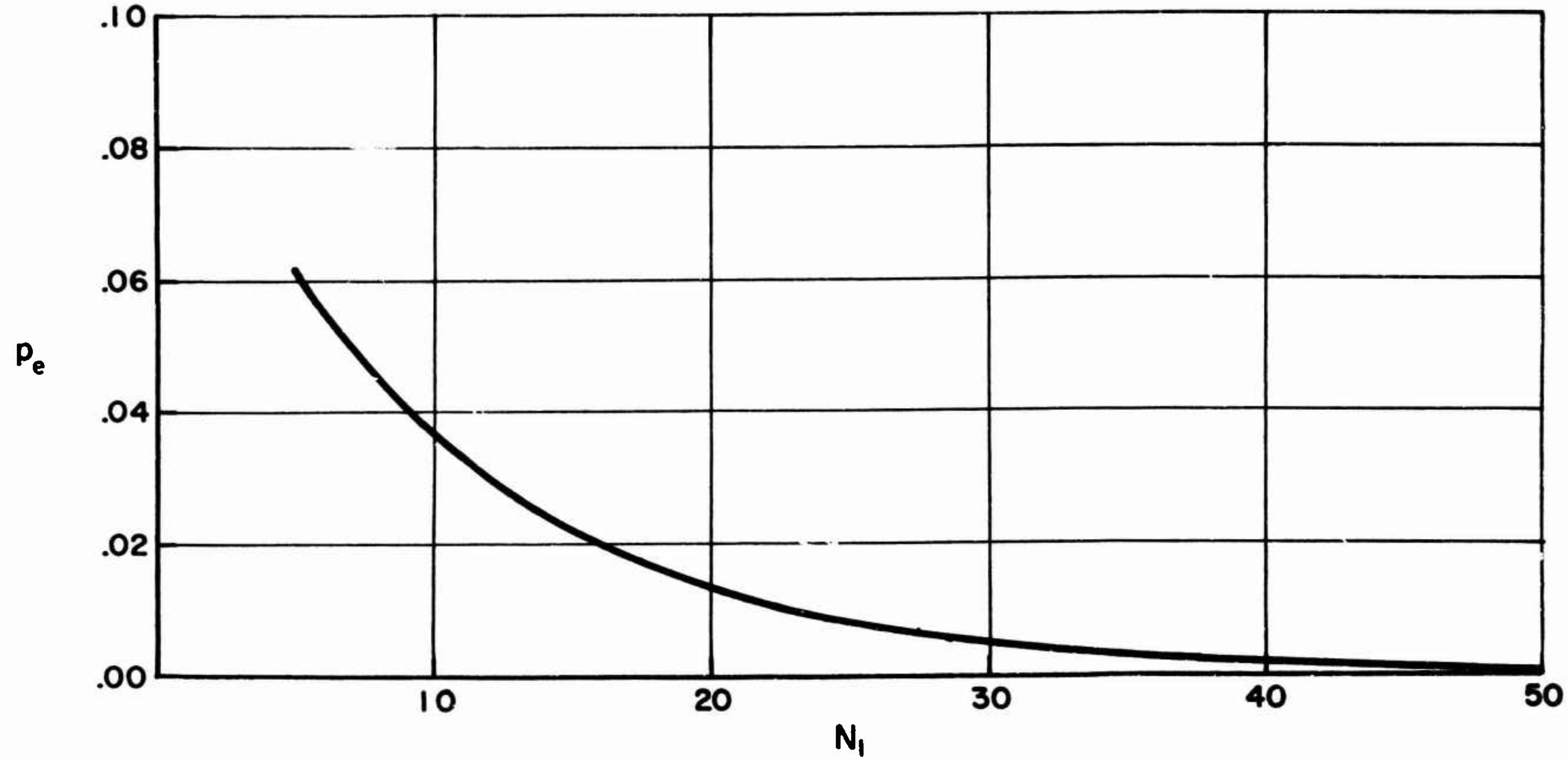

Fig. 12 GRAPH OF PROBABILITY OF ERROR AS A FUNCTION N_1 FOR $\beta = 0.9$, $\gamma_1 = 2.0$

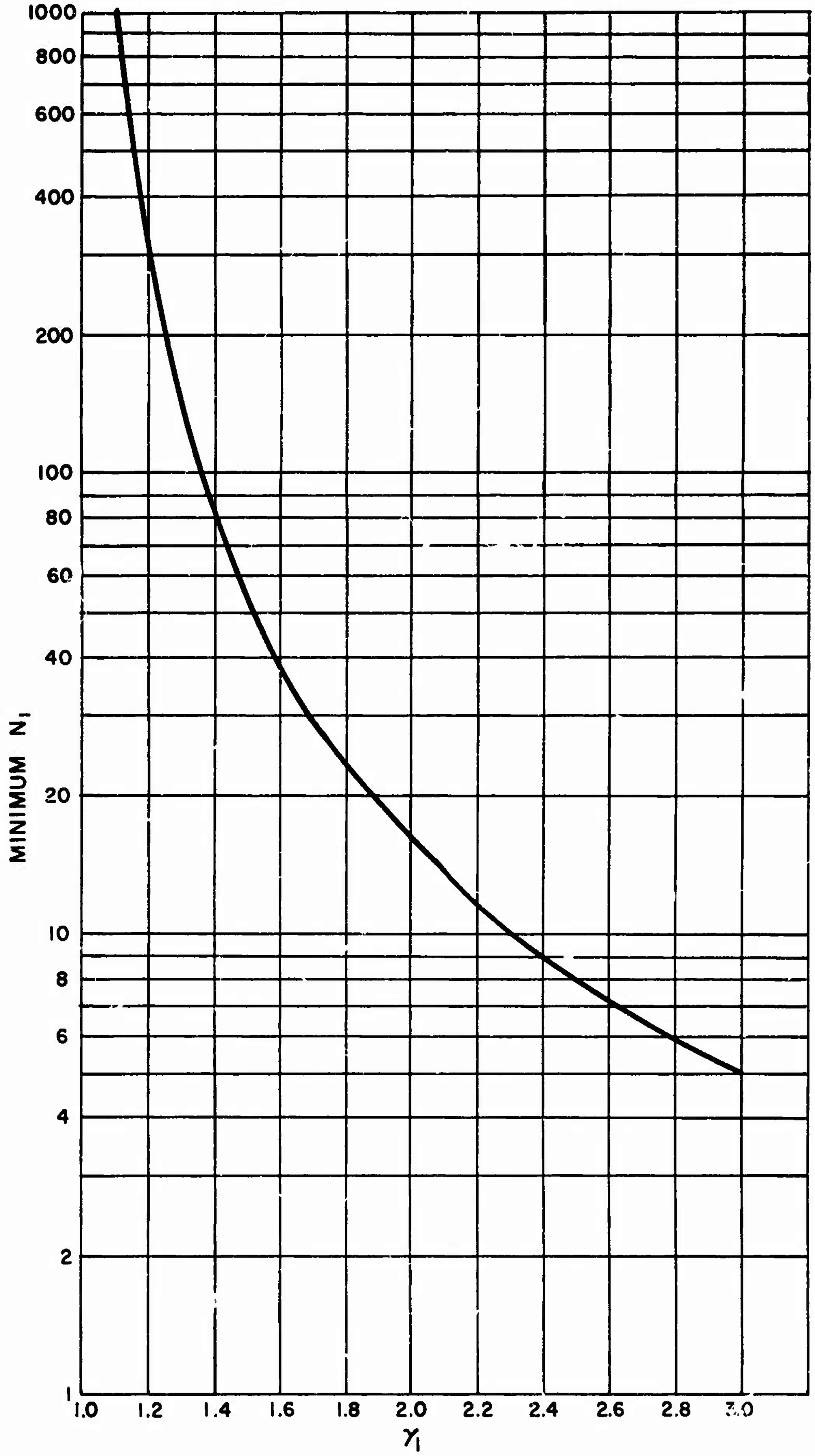

Fig. 13 MINIMUM EXPECTED NUMBER OF GRAINS REQUIRED IN TEST AREA TO SATISFY P = .05, Q = .05 FOR $\lambda_1 > 1$

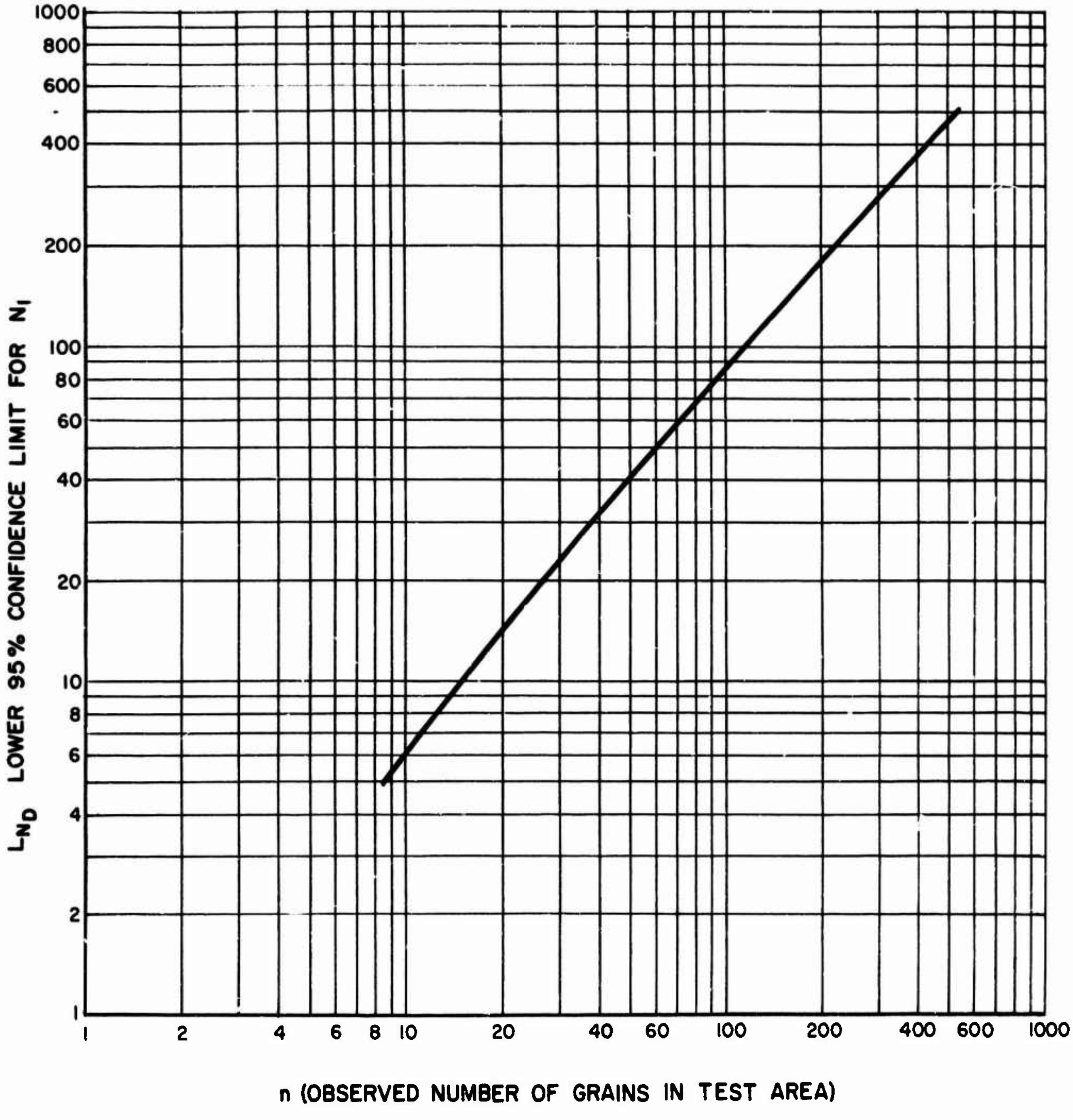

Fig. 14 LOWER 95% CONFIDENCE LIMIT FOR N_1 ($\gamma_1 > 1$)

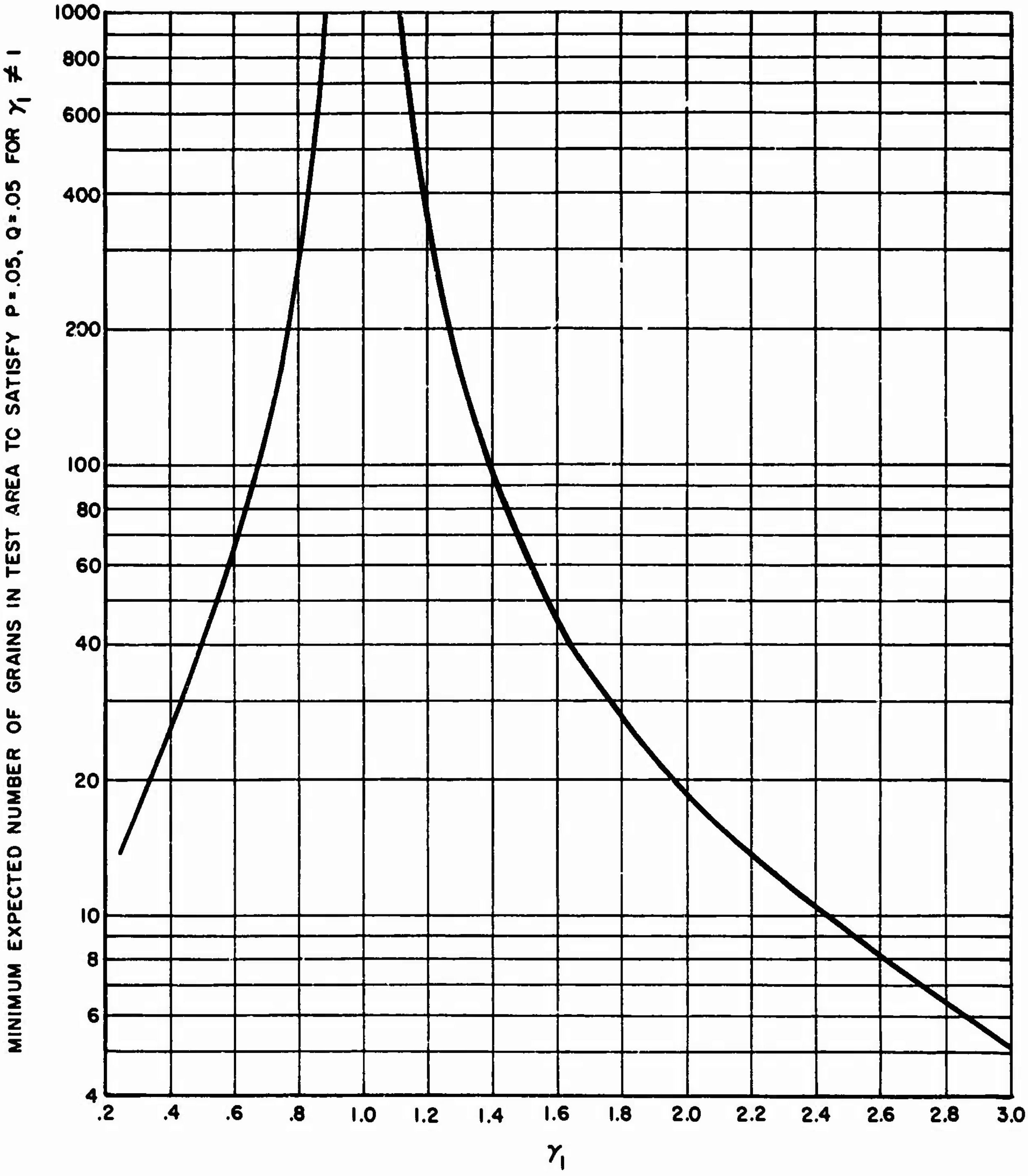

Fig. 15 MINIMUM EXPECTED NUMBER OF GRAINS REQUIRED IN TEST AREA TO SATISFY P=.05, Q=.05 FOR $\gamma_1 \neq 1$

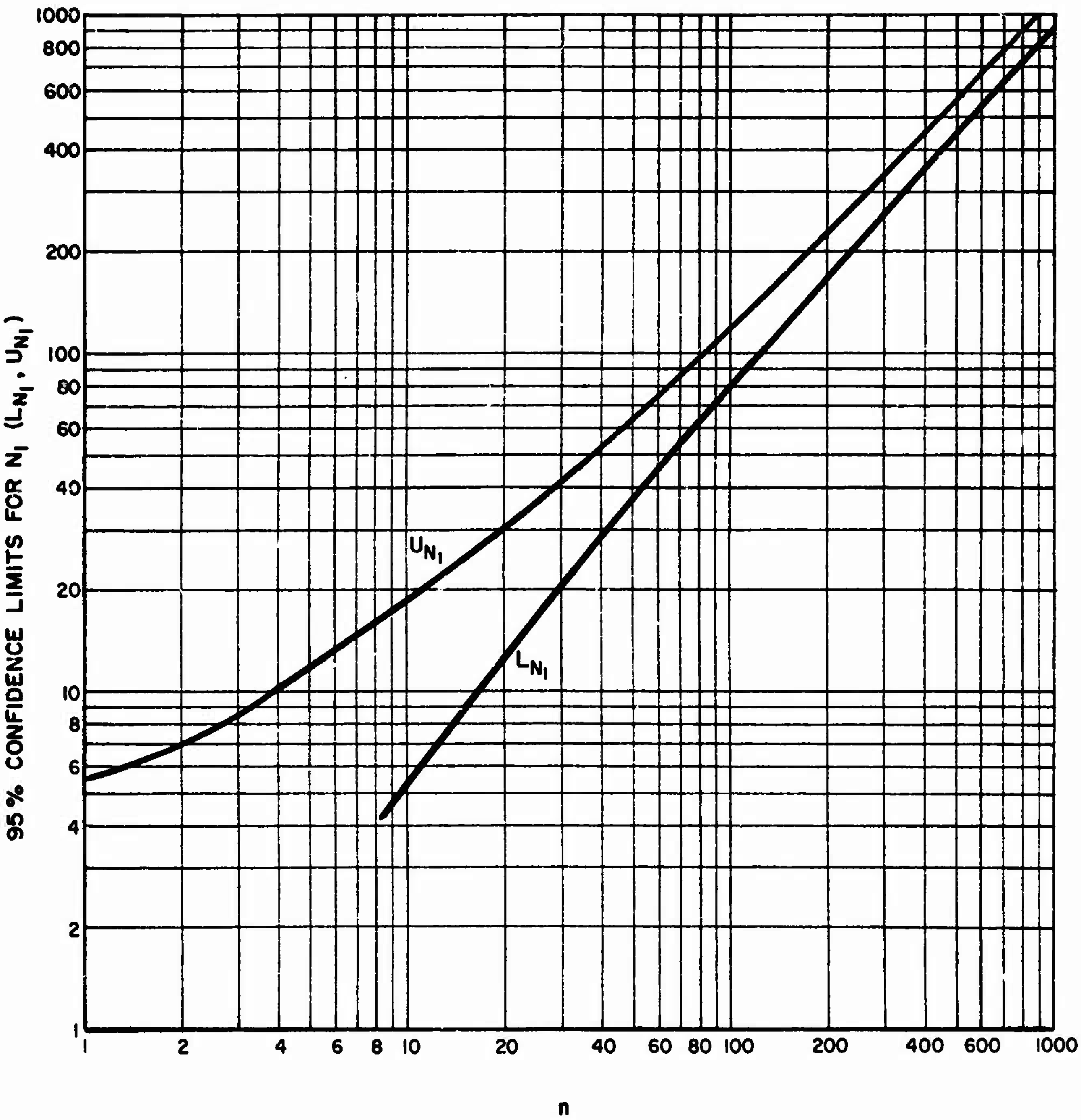

Fig. 16 95% CONFIDENCE INTERVALS FOR N_1 ($\gamma_1 \neq 1$)

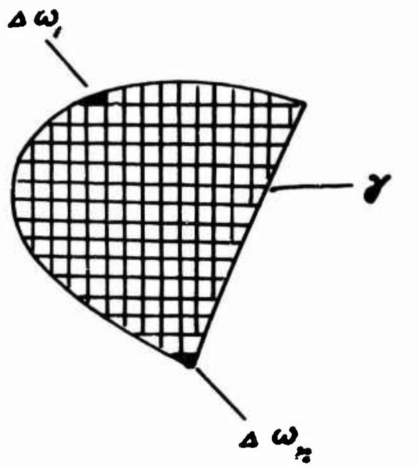

Fig. 17 SUBDIVISION OF THE AREA γ

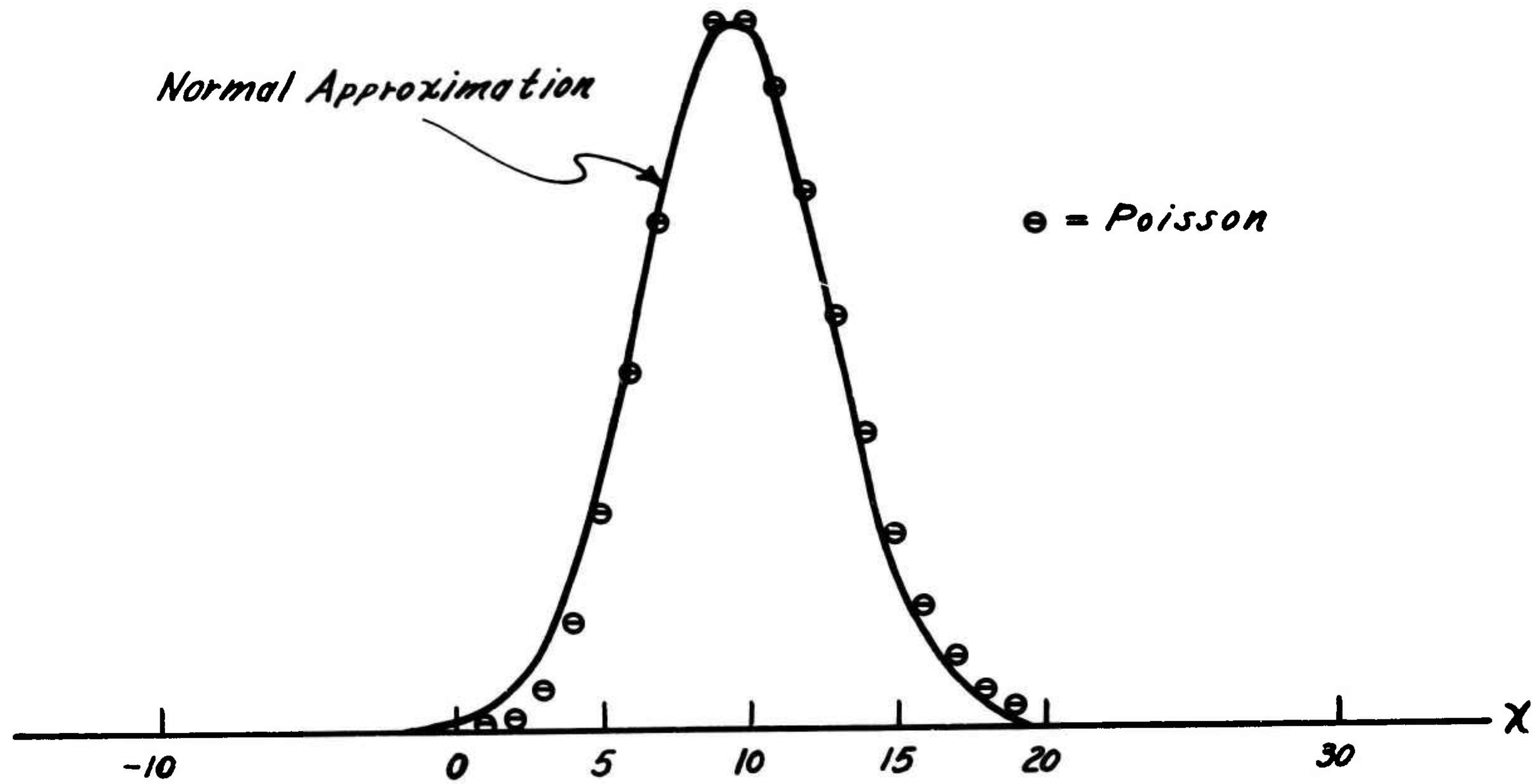

Fig. 18 NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION FOR N = 10

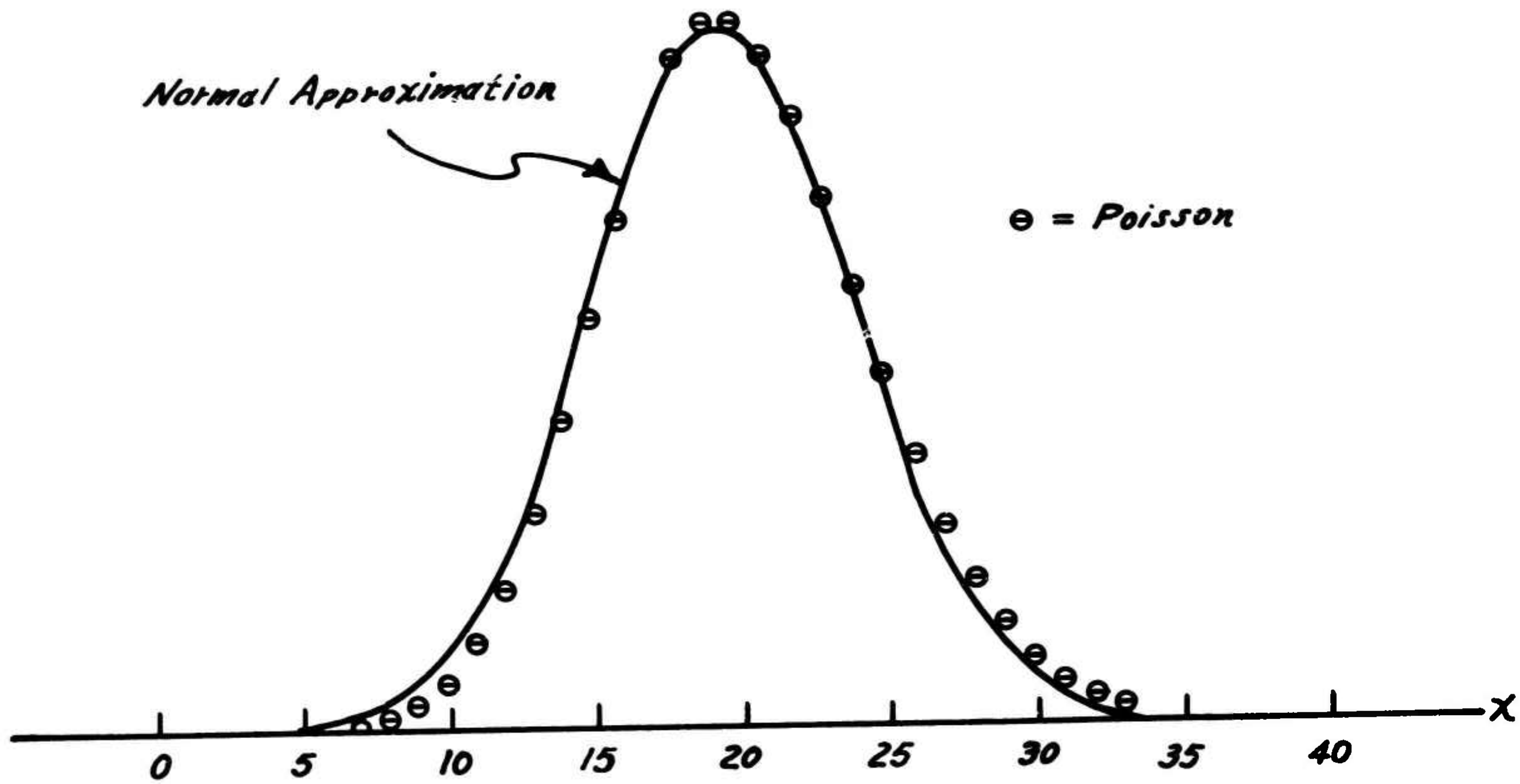

Fig. 19 NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION FOR N=20

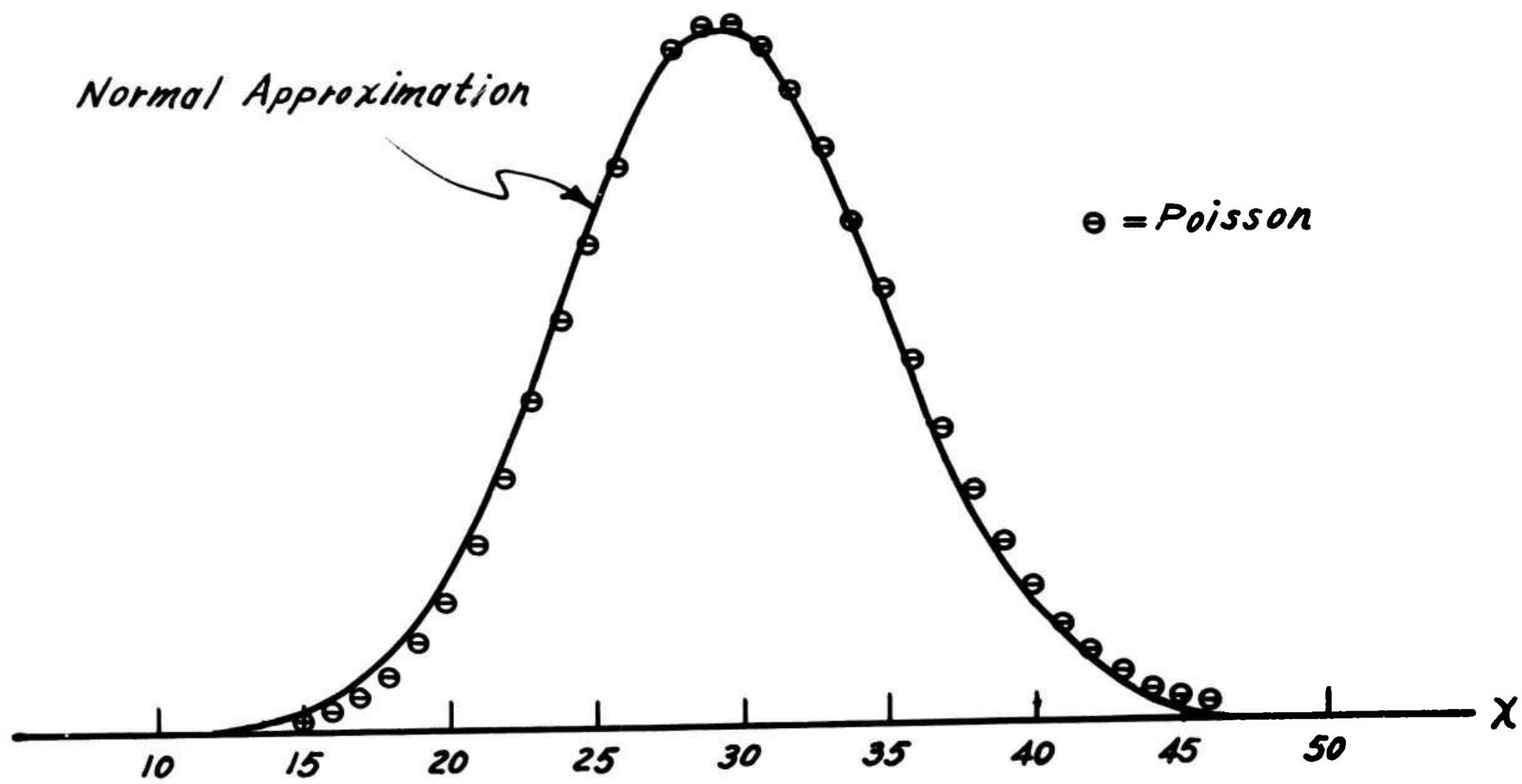

Fig. 20 NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION FOR N = 30

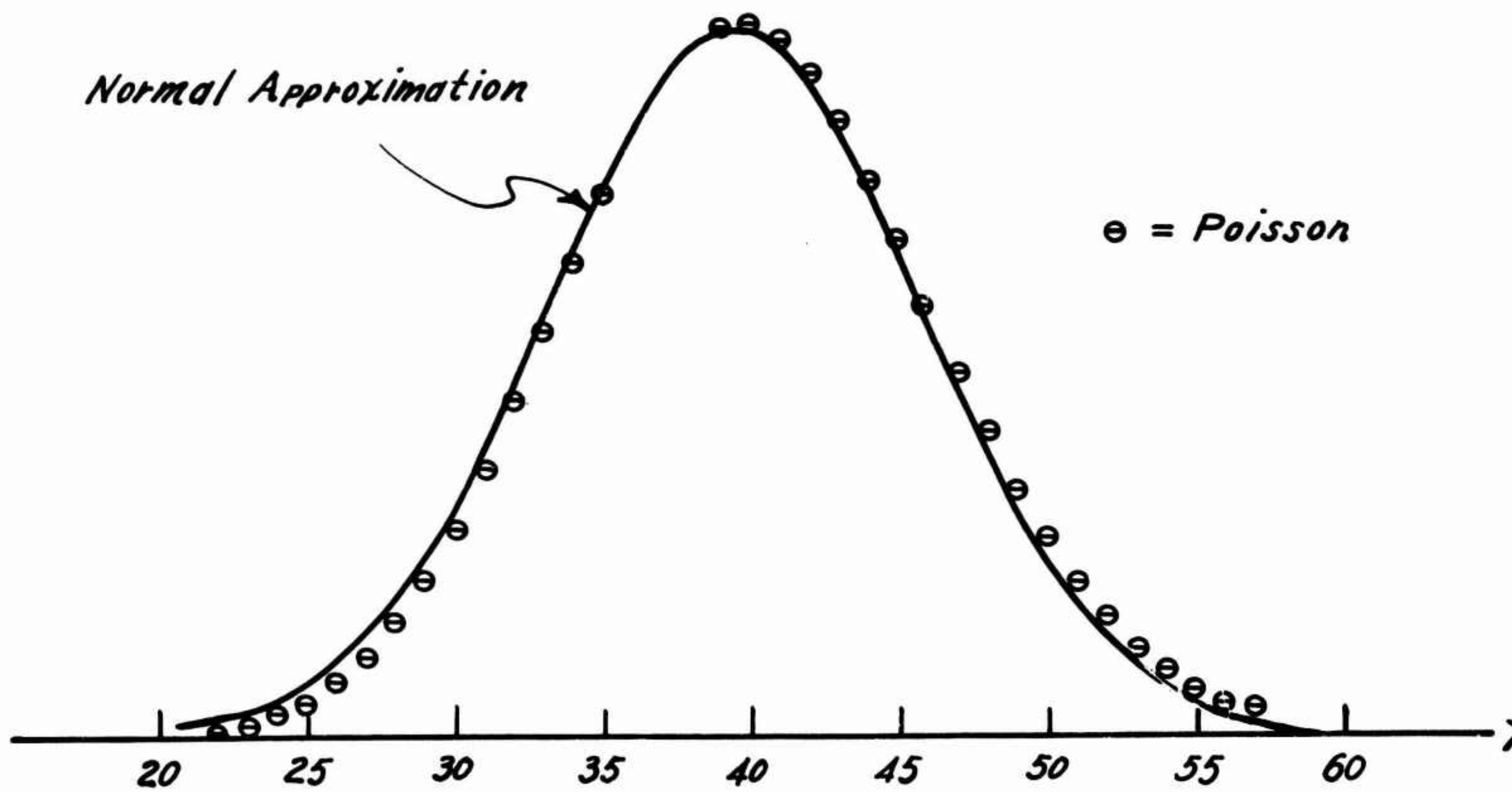

Fig. 21 NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION FOR N = 40

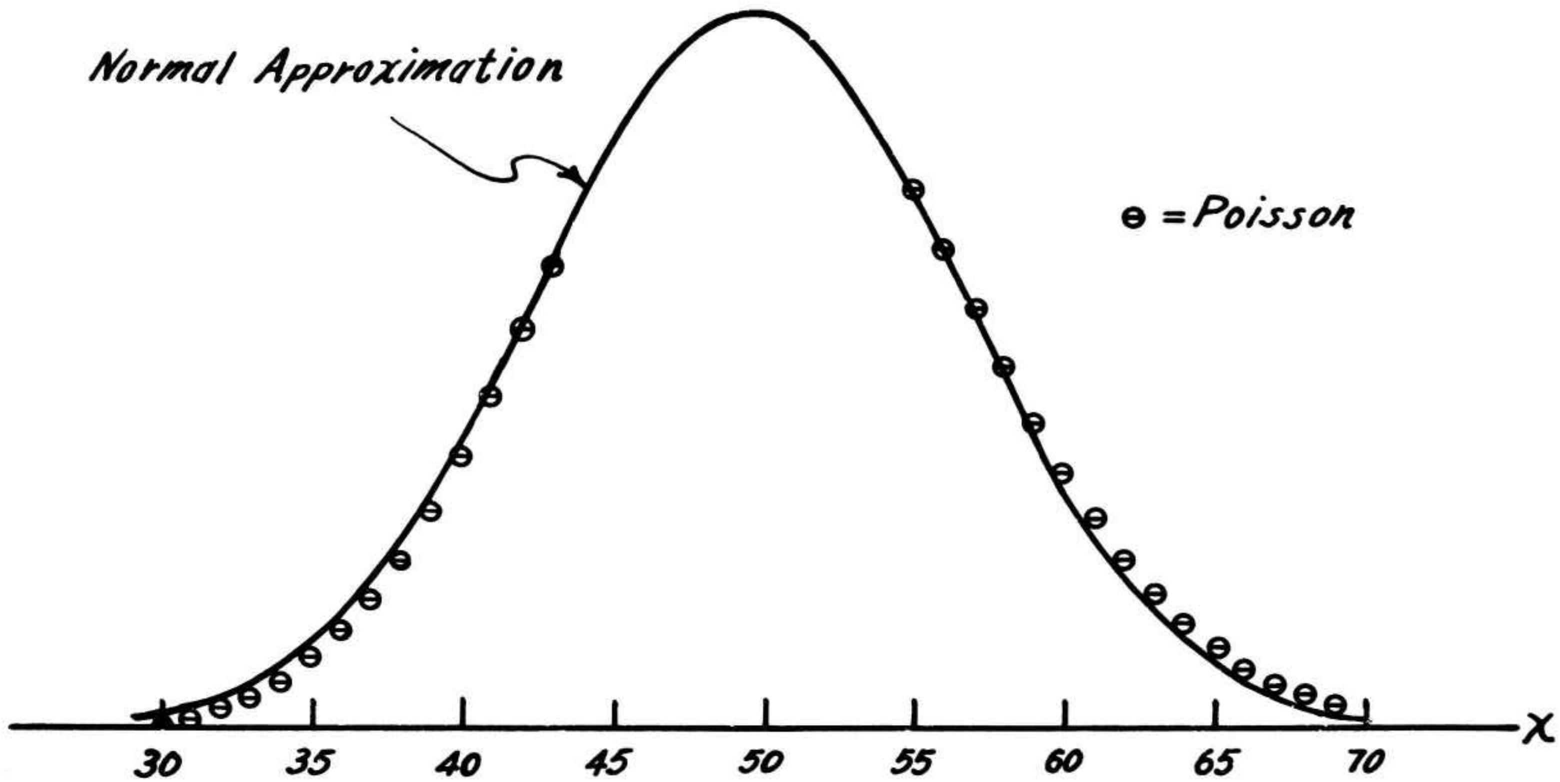

Fig. 22 NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION FOR N = 50

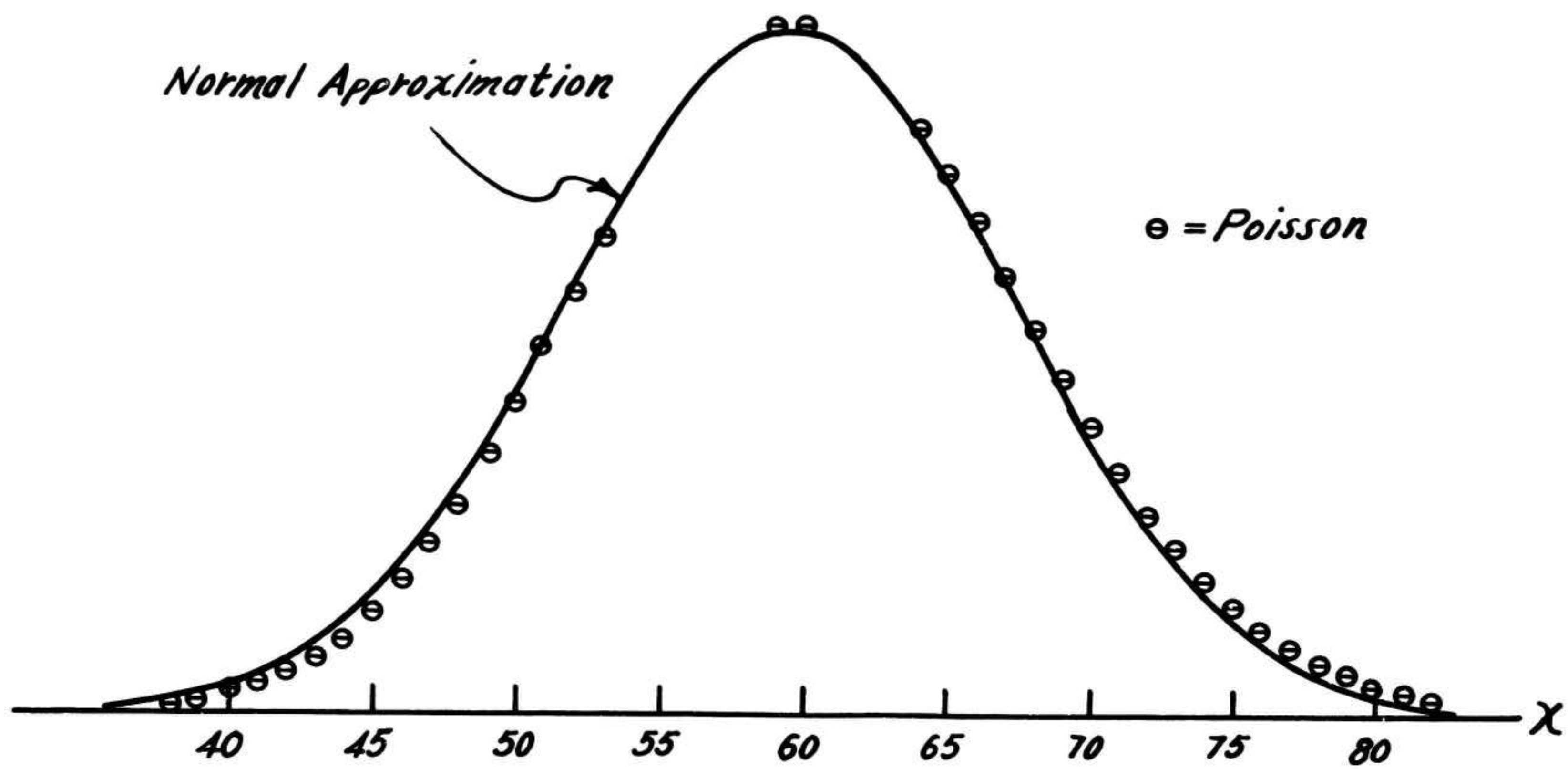

Fig. 23 NORMAL APPROXIMATION TO THE POISSON DISTRIBUTION FOR N = 60

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)

Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities

I. Guderley, K.G.

II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)

Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities

I. Guderley, K.G.

II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)
Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities

I. Guderley, K.G.
II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

Aeronautical Research Laboratory, Wright-Patterson Air Force Base, Ohio. ON THE EVALUATION OF STRONGLY ENLARGED PHOTOGRAPHS, by K.G. Guderley and M.D. Lum. February 1961. 103 p. incl. illus. tables. (Project 7071, Task 70437) (ARL TR 60-275)
Unclassified Report

The accuracy of the evaluation of a photographic plate is limited by its grain structure. One approximates the value for the light density at a given point by the average light density in a small area (the "test area") surrounding the point. This paper establishes confidence limits for evaluation procedures of this kind. It is assumed that the grains on the photographic plate arise in independent random processes controlled by the local density of the light flux. In

(over)

UNCLASSIFIED

1. Photographs - enlargements
2. Photographs - light densities

I. Guderley, K.G.
II. Lum, M.D.

UNCLASSIFIED

the evaluation procedure one counts the number of grains in the test area. Generalizing the method one attaches to each grain a weight factor depending upon the grain position within the test area and then determines the sum of the weight factors for the grains found in the test area. By such a procedure one can determine quantities related to the light density, e.g. the density gradient; one also can scan for patterns of a special kind, e.g. a sudden jump of the light density. For measurements of this kind probability theory predicts the expected value and the variance in terms of the light density and the chosen weight function. There are two kinds of errors due to the randomness inherent in the process of grain generation. The variance due to errors of both kinds must be minimized.

UNCLASSIFIED

UNCLASSIFIED

UNCLASSIFIED